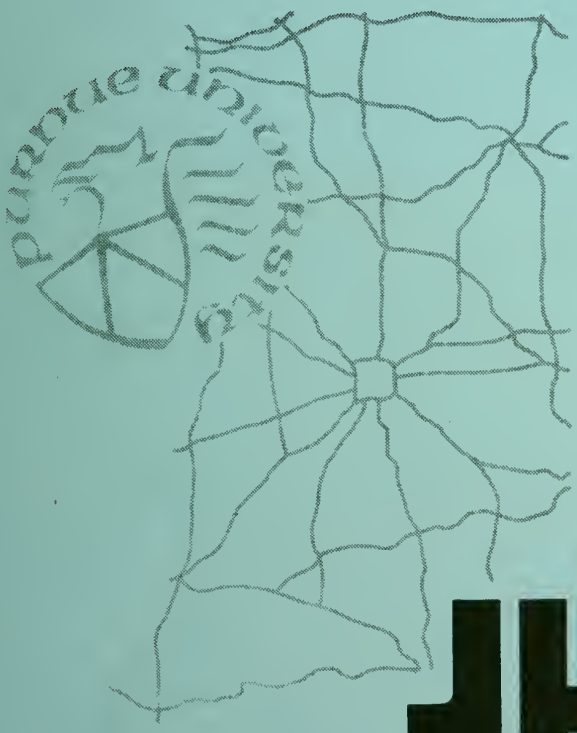
# MODELING THE IMPACT OF HIGHWAY IMPROVEMENTS ON THE VALUE OF ADJACENT LAND PARCELS

DECEMBER 1969 – NUMBER 37




BY

E. I. ISIBOR

JHRP

JOINT HIGHWAY RESEARCH PROJECT

PURDUE UNIVERSITY AND

INDIANA STATE HIGHWAY COMMISSION

Progress Report

# MODELING THE IMPACT OF HIGHWAY IMPROVEMENTS ON THE VALUE OF ADJACENT LAND PARCELS

TO: J. F. McLaughlin, Director
Joint Highway Research Project

December 30, 1969

File No: 3-5-7

FROM: H. L. Michael, Associate Director
Joint Highway Research Project

Project No: C-36-64G

Attached is a Progress Report on the HPR research project "Highway Impact Studies". This Report is titled "Modeling the Impact of Highway Improvements on the Value of Adjacent Land Parcels". It has been prepared by Mr. Edward I. Lubbor, Graduate Instructor in Research on our staff, under the direction of Professor H. L. Michael.

The research here reported was directed at developing techniques for predicting the impact of highway improvements on the value of adjacent land parcels. Two techniques are discussed. Analytical models were developed from Indiana and Florida data and important factors affecting land value were determined. A conceptual framework was also developed for a non-deterministic approach which shows promise and perhaps is more realistic for use.

The report is presented for the record and will also be submitted to the ISHC and the BPR for review, comment and acceptance.

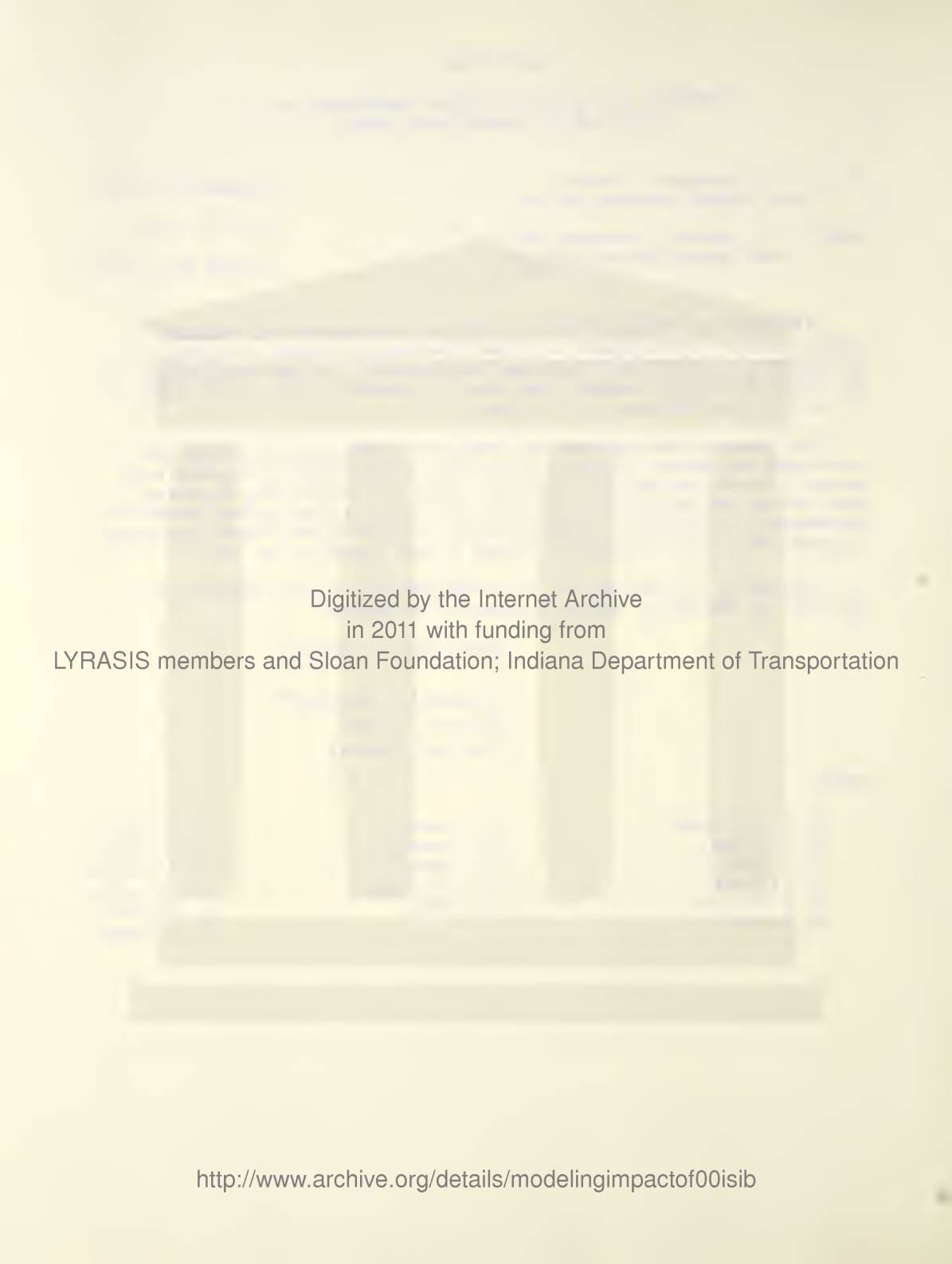
Respectfully submitted,

Harold L. Michael

Harold L. Michael
Associate Director

HLM/rg

cc: F. L. Ashbaucher
W. L. Dolch
W. H. Goetz
W. L. Grecco
G. K. Hallock
M. E. Harr
R. H. Harrell
J. A. Havers
V. E. Harvey
F. B. Mendenhall
R. D. Miles
C. F. Scholer
M. B. Scott
W. T. Spencer
H. R. J. Walsh
K. B. Woods
E. J. Yoder



Digitized by the Internet Archive
in 2011 with funding from
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

http://www.archive.org/details/modelingimpactof00isib

Progress Report

# MODELING THE IMPACT OF HIGHWAY IMPROVEMENTS ON THE VALUE OF ADJACENT LAND PARCELS

Edward L. Kaiboc
Graduate Instructor in Research

Joint Highway Research Project

Project C-36-64J
File 3-5-7

Prepared as Part of an Investigation

Conducted by

Joint Highway Research Project
Engineering Experiment Station
Purdue University

in cooperation with the

Indiana State Highway Commission

and the

U.S. Department of Transportation
Federal Highway Administration
Bureau of Public Roads

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Bureau of Public Roads

Not Released for Publication

Subject to Change

Not Reviewed by
Indiana State Highway Commission
or the
Bureau of Public Roads

Purdue University
Lafayette, Indiana
December 30, 1969

## ACKNOWLEDGEMENTS

The author is sincerely grateful to his Faculty Advisor, Professor Harold L. Michael for his assistance in the review of the manuscript and for his many valuable suggestions.

The encouragement and advising of Professors James R. Buck and Irving W. Burr on different aspects of this research are very much appreciated. Thanks is also due to Professors William L. Grecco and Joseph C. Oppenlander for their guidance during the conceptional phase of this work. Mr. Richard E. James is thanked for his assistance on various statistical topics.

The author is grateful to the Joint Highway Research Project of Purdue University, The Indiana State Highway Commission and the Bureau of Public Roads for both their cooperation and their financial support which made this research possible.

Special thanks to my dear wife, Edwina, for her encouragements throughout this work and for typing the thesis. The author is also grateful to his son, Ekinadose, for his patience during those long hours that he was not around to play with him.

TABLE OF CONTENTS


| | Page |
|---|---|
| LIST OF TABLES | vi |
| LIST OF FIGURES | viii |
| ABSTRACT | xi |
| CHAPTER ONE. INTRODUCTION | 1 |
| Transportation Planning as a Decision Process | 1 |
| The Impact Problem | 5 |
| Types of Highway Impacts | 8 |
| Aesthetic Impacts of Highway Transportation | 8 |
| The Social Impacts of Highway Improvement | 11 |
| The Political Impacts of Highway Improvement | 13 |
| The Economic Impacts of Highway Improvement | 13 |
| Measurement of Economic Impact | 16 |
| Purpose and Scope | 19 |
| CHAPTER TWO. A REVIEW OF THE STATE OF THE ART IN LAND ECONOMIC STUDIES | 22 |
| General | 22 |
| A Review of Highway Severance Damage Studies | 23 |
| Methodology in Severance Damage Studies | 27 |
| Findings from Severance Damage Studies | 28 |
| A Review of Economic Impact Studies | 35 |
| Techniques in Economic Impact Studies | 36 |
| Findings of Major Economic Impact Studies | 37 |
| Impact Research at Purdue University | 39 |
| CHAPTER THREE. AN 'ANATOMY' OF LAND VALUE CHANGES | 44 |
| Land Use | 46 |
| People's Attitude | 48 |
| Accessibility | 49 |
| The Distance of a Land Parcel from the Highway Improvement | 53 |

TABLE OF CONTENTS (continued)

Page

Distance from Nearest Urban Area ..... 56
Type of Area ..... 59
Time Elapsed ..... 59
Size of Land Parcel ..... 61
Type of Access Control ..... 61
Visual Exposure ..... 62

CHAPTER FOUR. DEVELOPMENT OF A DETERMINISTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM ..... 63

Modeling the Real World ..... 63
Regression Theory ..... 67
Calibration Procedure ..... 68
Statistical Measures ..... 71
The Coefficient of Multiple Determination ($R^2$) ..... 71
Partial F or Sequential F Test ..... 73
The t Statistic ..... 74
The Standard Error of Estimate ..... 74
Uses of Dummy Variables in Regression Analysis ..... 75
The Dummy Variable Technique ..... 77
First Type of Constraint ..... 78
Second Type of Constraint ..... 79
A Geometric Interpretation ..... 84
Measures of Importance ..... 86
Relative Importance of Each Dummy Variable Set ..... 87
Testing for Significance ..... 88

CHAPTER FIVE. DESIGN OF STUDY ..... 89

Methodology ..... 89
Variables Used in the Models ..... 92
Dependent Variables ..... 92
Independent Variables ..... 95
Data ..... 96
Source of Data ..... 96
Procedure for Obtaining Land Value Data ..... 97
Data Collected ..... 98

TABLE OF CONTENTS (continued)


| | Page |
|---|---|
| CHAPTER SIX. RESULTS OF MODEL DEVELOPMENT ..... | 108 |
| Indiana (I) Set of Models ..... | 108 |
| Indiana-Florida (IF) Set of Models ..... | 122 |
| CHAPTER SEVEN. DISCUSSION OF RESULTS ..... | 137 |
| Effect of Data Inadequacy on Results ..... | 137 |
| Comparison Between the I Set and the IF Set of Models ..... | 140 |
| Evaluation of Models in the I Set ..... | 142 |
| Relative Importance of the Independent Variables ..... | 154 |
| CHAPTER EIGHT. DEVELOPMENT OF A NON-DETERMINISTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM: A CONCEPTUAL FRAMEWORK ..... | 161 |
| The Impact Problem Revisited ..... | 161 |
| An Analogy Between the Impact Problem and Medical Diagnosis ..... | 163 |
| Outline of Bayesian Approach to the Impact Problem ..... | 166 |
| A Sample Problem ..... | 177 |
| Problem Statement ..... | 177 |
| Solution ..... | 177 |
| CHAPTER NINE. CONCLUSIONS ..... | 187 |
| CHAPTER TEN. RECOMMENDATIONS FOR EXTENSION AND FURTHER RESEARCH ..... | 191 |
| LIST OF REFERENCES ..... | 193 |
| VITA ..... | 204 |

# LIST OF TABLES


| Table | | Page |
|---|---|---|
| 1. | Mean Price Per Acre Along Lebanon Bypass | 54 |
| 2. | Mean Value Per Acre of Land Along Kokomo Bypass in Constant Dollars | 54 |
| 3. | Indiana Data | 100 |
| 4. | Florida Data | 102 |
| 5. | Summary of Model I-CCI | 109 |
| 6. | Summary of Model I-CC2 | 111 |
| 7. | Summary of Model I-CC3 | 112 |
| 8. | Summary of Model I-PCI | 114 |
| 9. | Summary of Model I-PC2 | 115 |
| 10. | Summary of Model I-PC3 | 117 |
| 11. | Summary of Model I-RR 1 | 118 |
| 12. | Summary of Model I-RR2 | 120 |
| 13. | Summary of Model I-RR3 | 121 |
| 14. | Summary of Model I-RR2A | 123 |
| 15. | Summary of Model IF-CCI | 124 |
| 16. | Summary of Model IF-CC2 | 125 |
| 17. | Summary of Model IF-CC3 | 127 |
| 18. | Summary of Model IF-PC 1 | 128 |
| 19. | Summary of Model IF-PC 2 | 129 |
| 20. | Summary of Model IF-PC3 | 131 |

| Table | | Page |
|---|---|---|
| 21. | Summary of Model IF-RR1 ..... | 132 |
| 22. | Summary of Model IF-RR2 ..... | 133 |
| 23. | Summary of Model IF-RR3 ..... | 135 |
| 24. | Summary of Model IF-RR2A ..... | 136 |
| 25. | Comparative Statistics for all Developed Models ..... | 141 |
| 26. | Values of Beta Coefficients for Variables in Each Model ..... | 155 |

# LIST OF FIGURES


| Figure | | Page |
|---|---|---|
| 1. | General Description of the Transportation Planning Process | 6 |
| 2. | Types of Highway Impacts | 9 |
| 3. | Victims and Beneficiaries of Highway Improvements | 17 |
| 4. | States Conducting Severance Damage Studies | 26 |
| 5. | Land Value Recovery Rates (Over-All), by Number and Percent of Cases | 29 |
| 6. | Land Value Recovery Rates, by Time From Acquisition to Sale, Unadjusted for General Land Value Changes | 29 |
| 7. | Land Value Recovery Rates, by Land Use at Time of Acquisition | 31 |
| 8. | Land Value Recovery Rates, by Type of Highway System | 31 |
| 9. | Land Value Recovery Rates, by Travel Distance to New Highway | 34 |
| 10. | Land Value Recovery Rates, by Nearness to Interchange (Over-All) | 34 |
| 11. | Location of Study Facilities in Indiana | 43 |
| 12. | General Price Trends | 47 |
| 13. | Opinions of Residents of Nearby Highway in Westchester County, New York (Traffic Impact) | 50 |
| 14. | Mean Land Value (Dollars/Acre) in Constant Dollars at Various Distances from the Improvement | 55 |
| 15. | Mean Land Value in Constant Dollars at Various Distances from the Improvement | 57 |
| 16. | Land Value Bands Parallel to the By-Pass | 58 |

| Figure | | Page |
|---|---|---|
| 17. | Land Price Changes Along Ventura Boulevard in Camarillo, California | 60 |
| 18. | The Role of Models in Analysis | 65 |
| 19. | Two Dimensional Regression | 72 |
| 20. | Effect of Membership in a Dummy Variable Class on Regression Line | 85 |
| 21. | Identification System for Developed Models | 90 |
| 22. | Mean Change in Land Values for Classes Within Each Variable | 138 |
| 23. | Mean Change in Land Values for Classes Within Each Variable | 139 |
| 24. | Comparison of Mean Recovery Rates for Parcels in Florida and Indiana | 143 |
| 25. | Plot of Observed Against Predicted Values of the Response Variable for Model I-CCI | 147 |
| 26. | Plot of Observed Against Predicted Values of the Response Variable for Model I-CC2 | 148 |
| 27. | Plot of Observed Against Predicted Values of the Response Variable for Model I-CC3 | 149 |
| 28. | Plot of Observed Against Predicted Values of the Response Variable for Model I-PC2 | 150 |
| 29. | Plot of Observed Against Predicted Values of the Response Variable for Model I-RR2 | 151 |
| 30. | Plot of Observed Against Predicted Values of the Response Variable for Model I-RR2A | 152 |
| 31. | Mean Change in Land Values for Classes Within Each Variable | 158 |
| 32. | Main Components of a Diagnostic Approach to the Impact Predictive Problem | 171 |

Figure Page

33. Stages in the Development of a Non-deterministic Approach to the Impact Predictive Problem ..... 178

34. Location of Parcels in the Sample Problem in Relation to the Highway Improvement ..... 179

35. Relating Descriptors of Land Parcels to Impact States .... 184

# ABSTRACT

Isibor, Edward Iroguehi, Ph.D., Purdue University, January, 1970. MODELING THE IMPACT OF HIGHWAY IMPROVEMENTS ON THE VALUE OF ADJACENT LAND PARCELS. Major Professor: Harold L. Michael.

Decisions on transportation improvements have ramifications that extend far beyond the system itself. A systematic evaluation of transportation alternatives therefore demands that a scrutiny of the total system effects be made if a reliable estimate of the benefit-cost relationship for each alternative is to be determined. Numerical measures are presently available for predicting the impacts of highway improvements on their users; but the evaluation of the impacts on non-users is still limited to identifying these impacts only in a descriptive "what is" or "what has been" context. This research was directed at developing techniques for predicting the impact of highway improvements on the value of adjacent land parcels.

Using data from the states of Indiana and Florida, two sets of regression models (the I and IF sets) were developed. The I set used only Indiana data while the IF set used the combined data from the two states.

Models developed in the IF set were rather weak in explaining the variation in land values resulting from the highway improvement. On the other hand, models developed in the I set performed very well in this regard. Variables defining the type of highway, the type of land

use before the highway improvement, the type of access control between a land parcel and the highway improvement, the areal location of the parcel, the size of the parcel and the time elapsed after completion of the highway improvement explained to a substantial degree (slightly more than 87% in one of the developed models) the variation in mean change of the value of Indiana parcels following a highway improvement. Further analysis of the limited Indiana data also revealed that variables denoting highway type areal location of the parcel and land use type were most important in explaining the observed changes.

The use of dummy variables proved very effective for including discontinuous variables in the developed regression models.

This research also constructed a conceptual framework for developing a non-deterministic approach to the impact predictive problem. In this approach the impact problem was conceptualized as being identical to the medical diagnosis problem. Building on this analogy, the impact problem then becomes that of predicting the probable impact states of a land parcel following a highway improvement on the basis of some observable or measurable descriptors. Data from past occurrences of highway impacts are therefore needed to generate a priori distributions about the probable occurrence of various impact states. An application of Bayes' theorem of conditional probability then provides a linkage between the a priori probability distribution and the required posterior probabilities. Past data might not always be available or might be too expensive to obtain. In such cases, the decision maker might have to rely on the subjective evaluations of a team of experts. A single probability distribution representing a consensus of the experts'

evaluations is then obtained for a later use in a formal Bayesian analysis. This approach appears to be more realistic and does help in revealing a fuller spectrum of the impact states that might occur following a highway improvement than the one-shot estimate provided by regression equations.

The results of this research recommend that further studies be undertaken with a much larger data base to verify the conclusions reached in this work. A study should also be made to fully develop the non-deterministic approach suggested in this work.

# CHAPTER ONE

# INTRODUCTION

Transportation systems have played an indispensable role in the development of our civilization. They have enabled our technology to keep pace with the rising tide of man's expectations and needs. They have helped to stimulate interactions between man and his environment and they have served as an impetus to economic activity. This implies that a "cause and effect" relationship can be associated with a transportation system. This relationship was stressed by the late Thomas H. McDonald of the Bureau of Public Roads in these words:

> "We were not a wealthy nation when we began improving our highways...but the roads themselves helped us create a new wealth in business and industry and land values... So it was not our wealth that made our highways possible. Rather it was our highways that made our wealth possible." (1)

Indeed transportation improvements have become one of the greatest instigators of change in our economic, social and political institutions. However, changes occur continuously in the general physical world; no material thing remains static. But transportation improvements by their nature act as catalysts which speed up this rate of change. This change which a highway improvement induces on its surrounding environment is often referred to as its "impact."

## Transportation Planning as a Decision Process

In studying the impact of various transportation alternatives it is informative to define what we understand by transportation itself.

Transportation may be defined as the translocation of matter in physical space both in time and in state. This translocation of matter helps to enhance its value. Objects, be they persons or goods, have little value unless given utility, the capacity for being useful or for satisfying wants.

In agreement with the above concept on transportation, Hay (2) asserts that transportation contributes two kinds of utility to goods - place utility and time utility. Simply stated, this means having goods where they are wanted when they are wanted.

Transportation therefore is nothing more than a change in the state of objects in which each object is moved in time and physical space and in which other attributes of each object are also likely to undergo a change. Certain motivational forces operate within our physical space to determine the trip vector (defined to mean the length and direction from origin to destination). These motivational forces could be the desire to go to the swimming pool, to the grocery, to the place of work or the need to deploy troops to strategic locations in the event of war. Thus we find that transportation is hardly ever desired for its own sake; it merely enables us to satisfy other objectives. We can therefore, with ample justification, label our demand for transportation as a derived demand.

To meet this demand an adequate transportation system is required. A transportation system is referred to here to mean the totality of those physical components, rules or procedures of operation which are employed in the satisfaction of transportation needs.

The transportation problem which such a system is designed to

solve has many recognizable facets. Neufville (3) explained it this way:

"In brief the transportation problem is:

- multivariate: a large number of parameters is associated with even the simplest transportation problem
- stochastic: the demand for transportation services varies randomly over time as well as over space
- periodic: the use of a system fluctuates in cycles daily, weekly and seasonally
- non-stationary: the variations in demand do not fluctuate about a constant mean but rather are distributed about trends which may themselves be non-linear".

The transportation planning process serves the intermediary role of transforming expressed transportation needs into a transportation system which meets those needs, subject to certain constraints. Such constraints include the following:

Technology:

The present level of technology delimits the space in which feasible solutions are possible.

Monetary:

An acceptable solution must be within the financial resources of those for whom the system is to be built.

Resources:

A feasible solution must not require more resources than the society is willing to apportion to the transport sector of the economy. This constraint is significant in that it focuses

the planners attention to the fact that transportation objectives form only one subset of the total set of goals in urban life.

Political:

The established policies governing financing, operation of the system and the political goals of the area must not be violated by any solution. Constraints emerging from zoning policies also fall under this category.

Physical:

Natural barriers such as a body of water, mountain ranges and man-made barriers might confront the planner with problems that he can only deal with at prohibitive costs.

The problem facing the transportation planner then reduces to finding the best combination of resources satisfying the above constraints.

With such a wide spectrum of options before him, the planner must have a set of criteria to provide him with a rational basis for decision making. Such decision making in transportation planning is made more complex by the fact that decisions made with respect to a transport system have ramifications that extend far beyond the bounds of the specific system. The transportation planner must therefore consider in a systematic manner the consequences in all their multiplicity before choosing among alternative systems. The degree of success with which he can screen out the best alternative from such a set depends to a large extent on how knowledgeable he is about the probable impacts of each plan. Unfortunately however, the complexity of the problem is

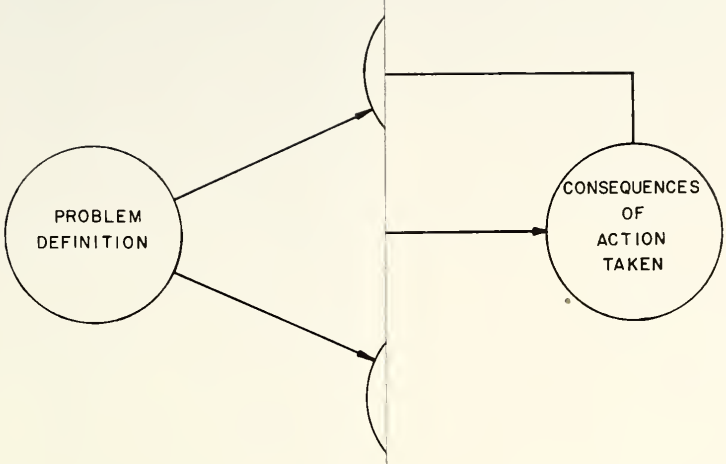
such that the planner cannot precisely prescribe what the actual outcomes will be when he makes his decision. This is so because the subsequent outcome depends on a number of different combinations of events, and at the time the decisions are made the decision maker is uncertain about what set of events will occur. All this points to the inevitable conclusion that every decision which a transportation planner makes involves some element of uncertainty. In this type of a situation it is essential that from time to time one looks back at past decisions, and evaluates the expected results with what actually occurred.

This feedback as suggested by Figure 1 is consistent with the general concept of planning. Planning is a continuous process and necessitates the constant reexamination of trends, tendencies, and policies in order to adapt and adjust governmental policies with the least possible friction and loss. Planning is thus an intellectual process, the conscious determination of courses of action, the basing of decisions on purpose, facts and considered estimates.

Impact studies are designed to supply facts about past decisions so as to facilitate decision-making in the future.

## The Impact Problem

In recognition of the useful knowledge that can be gained from impact studies, the American Association of State Highway Officials passed a resolution at their annual business meeting in November 1956 urging all state highway departments to promote research into the economic impact of highway improvements. In addition, they called upon the Highway Research Board to sponsor a conference with the sole




LEVELS IN PLANNING HIERARCHY

1

THE TRANSPORTATION PROBLEM THAT NEEDS TO BE SOLVED MUST BE PRECISELY DESCRIBED AT THIS STAGE. THIS STAGE IS VERY IMPORTANT INDEED IF A PROBLEM IS WELL DEFINED THE TASK UP — STAGE IN THE HIERARCHY IS GREATLY FACILITATED.

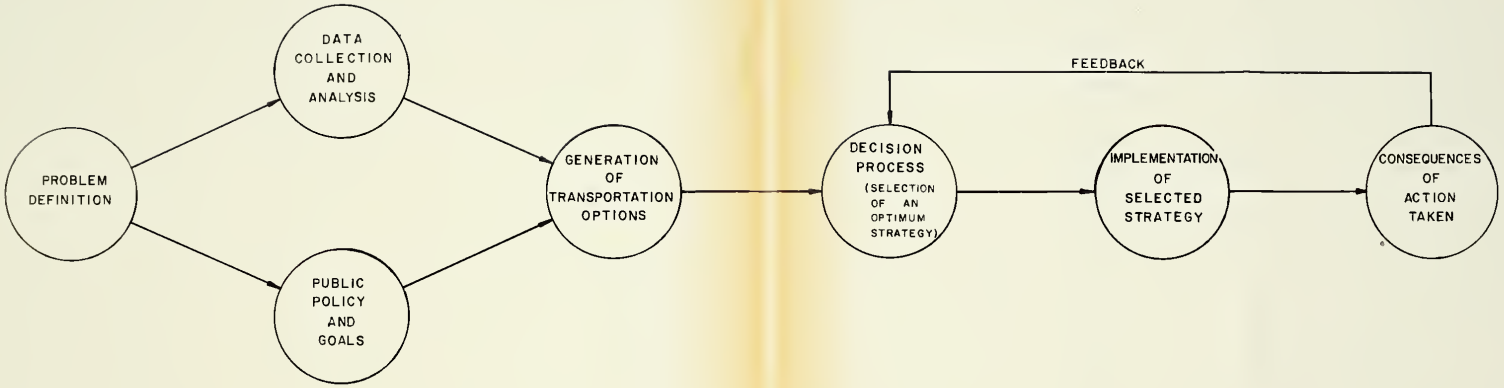
DATA TED
ARE COL AT
GET AN G
OF THE R
INFORMATI
AND ON
BE USED

PROJ
STATE O
AND OF
RESOURCES

6

THIS STAGE OBSERVES THE CONSE — QUENCES THAT OCCUR WHEN THE SELECTED STRATEGY WAS IMPLEMENTED IN THE REAL WORLD. ENOUGH TIME MUST BE ALLOWED HERE TO OBSERVE THE 'TOTAL' EFFECT SINCE SOME OF THEM ARE CUMULATIVE IN NATURE.

A FEEDBACK IS NECESSARY FROM THIS STAGE TO LEVEL 4 BECAUSE WHEN THE DECISIONS WERE MADE THE PLANNER WAS UNCERTAIN AS TO THE EVENTUAL OUTCOME. THIS FEEDBACK IMPROVES HIS DECISION MAKING IN THE FUTURE.




| LEVELS IN PLANNING HIERARCHY | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | THE TRANSPORTATION PROBLEM THAT NEEDS TO BE SOLVED MUST BE PRECISELY DESCRIBED AT THIS STAGE. THIS STAGE IS VERY IMPORTANT. INDEED IF A PROBLEM IS WELL DEFINED THE TASK UP – STAGE IN THE HIERARCHY IS GREATLY FACILITATED. | DATA PERTINENT TO THE PROBLEM ARE COLLECTED AT THIS STAGE TO GET AN INSIGHT INTO THE DEPTH OF THE PROBLEM AND TO GET INFORMATION ON AVAILABLE RESOURCES AND ON PUBLIC POLICIES WHICH MUST BE USED AS GUIDELINES. PROJECTION INTO THE FUTURE STATE OF THE PROBLEM AT HAND AND OF THE FINANCIAL AND OTHER RESOURCES NEEDED ARE MADE HERE. | OPTIONS ARE GENERATED AT THIS LEVEL WHICH SOLVE THE PROBLEM DEFINED IN LEVEL 1 AND WHICH SATISFY BOTH THE RESOURCE AND PUBLIC POLICY CONSTRAINTS IN LEVEL 2. | AT THE DECISION PHASE THE ALTERNATIVES GENERATED AT LEVEL 3 ARE COMPARED ON THE BASIS OF PREDETERMINED CRITERIA. INCLUDED HERE ARE PREDICTIONS OF THE PROBABLE IMPACTS OF EACH ALTERNATIVE ON ALL COMPONENTS OF THE SYSTEM. THAT ALTERNATIVE WHICH BEST MEETS THESE CRITERIA IS SELECTED. | THE ALTERNATIVE SELECTED IN LEVEL 4 IS IMPLEMENTED AT THIS STAGE THUS TRANSFERING PLANS DEVELOPED ON PAPER INTO ACTION IN THE REAL WORLD. | THIS STAGE OBSERVES THE CONSE – QUENCES THAT OCCUR WHEN THE SELECTED STRATEGY WAS IMPLEMENTED IN THE REAL WORLD. ENOUGH TIME MUST BE ALLOWED HERE TO OBSERVE THE 'TOTAL' EFFECT SINCE SOME OF THEM ARE CUMULATIVE IN NATURE. A FEEDBACK IS NECESSARY FROM THIS STAGE TO LEVEL 4 BECAUSE WHEN THE DECISIONS WERE MADE THE PLANNER WAS UNCERTAIN AS TO THE EVENTUAL OUTCOME. THIS FEEDBACK IMPROVES HIS DECISION MAKING IN THE FUTURE. |

FIGURE 1 GENERAL DESCRIPTION OF THE TRANSPORTATION PLANNING PROCESS

objective of discussing this matter.

The Highway Research Board, in accepting this recommendation, assembled, for the first time in the United States, most of the noted experts in the area of economic impact of highways. The momentum generated at this conference sparked much interest in this area among several state highway departments. By 1960, one hundred studies had been reported and more than forty additional investigations in twenty-nine states were under way (4). In these investigations, the Bureau of Public Roads and the state highway departments, working closely with many universities, attempted to delineate the various impacts accruing to communities from different types of highway improvements.

Most of these studies emphasized only the economic aspect of the total observed impact. However if we adhere to the definition of impact as the total change which a given highway improvement induces on all facets of the surrounding environment, then at least four impacts - aesthetic, social, political and economic - must be considered in an overall appraisal of impact. Clearly, all these are important. It can logically be asserted, however, that the aesthetic impact of a highway improvement, for instance, reveals itself partly as either an economic or social impact. In other words the impact of a highway on those using it or those who inhabit its environment has an ultimate effect on the way in which these people are impelled to order their economic and social activities and the degree of satisfaction derived therefrom. Similarly one can argue that social impacts of highway improvements are in part manifested through economic measures. People in a community have a discernible preference for one alternative social situation as compared to another. This preference is consciously or

unconsciously reflected in how much they are willing to pay for it. The foregoing is not to negate the need for studying other forms of impact apart from the economic aspect. Rather this is intended to point out that some correlation does exist between these four types of impact. This overlapping that exists among the types of highway impacts is depicted by Figure 2. Before we deal exclusively with the economic aspect of highway impact in this research, perhaps we should examine the various types of highway impact.

## Types of Highway Impacts

### A. Aesthetic Impacts of Highway Improvements

Great interest in the quality of the environment is unmistakably evident among the American people. This concern of citizens about the aesthetics of their environment has greatly changed the design philosophies of highway engineers. Design must now be compatible with the environment and must enhance rather than depreciate its aesthetic value. Today

> "more and more people are concerned with water pollution, air pollution ..... solid wastes, the preservation of areas for outdoor recreation and for open space, the design and arrangement of both the urban and rural landscape ..... living in harmony with nature has become a matter of conscious attention and national policy ....." (5)

The Highway Beautification Act of 1965 has added more urgency to aesthetic considerations in highway design. The impact of the highway beautification program is not susceptible to monetary quantification. This difficulty was encountered in a study of the economic impact of

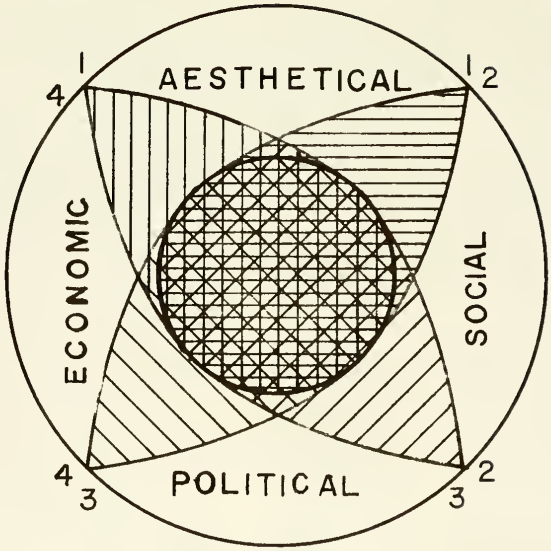



FIGURE 2: TYPES OF HIGHWAY IMPACTS

the Highway Beautification Act which the Bureau of Public Roads conducted in cooperation with several state highway department. Difficulty in measuring aesthetic values, however, does not negate their importance to the public. Scenic highways provide benefits at least to two groups; highway users and nearby landowners. Various surveys of motorists' preferences reveal that scenery is a very significant factor. In fact, in one report scenery was rated more important than travel time but less important than congestion. The longer the journey, the greater the probability that the driver will choose a scenic route (6). Cherner in his investigation of the influence of highway landscape development on property values remarked:

> "On Kennedy Expressway the apartment-dwellers and home-owners were most happy to exchange a crowded built-up street location for the open-cut area of greenery which this beautiful highway presented. This green belt in fact drew people to its borders for the open beauty as well as for the convenience of transportation.
>
> "On the Eisenhower Expressway in Chicago the road was a definite factor in revitalizing the slum areas adjacent to the highway. Obviously the grass and plantings encouraged these people to "breathe" again and certainly stimulated the cleaning and rehabilitation of the area. In the Maywood section of the Eisenhower Expressway, the residents have come to regard this area as their park to be enjoyed visually not physically. Many of these residents were relocated from their original homes on the right of way preferred to remain in the neighborhood and were happy with their decision because of the beauty of the expressway planting" (7).

These comments were based on a house-to-house survey conducted to sample people's view. This survey clearly indicated the pleasure that scenic highways or parkways bring to highway users and landowners. This pleasure may be reflected in a general way in land value.

B. The Social Impacts of Highway Improvements

The social impacts of highway improvements may be viewed as those effects which alter the relationship between people and social institutions such as the family, schools, churches and government. Attempts have been made by many studies to solicit from people themselves their views about these changes. Barbara Kemp (8) in her investigation of the social impact of a highway on an urban community noted that:

1. Most residents interviewed did not want to be dislocated. Most of them felt some intense attachment to their neighborhood.
2. Most argued that no displacements should occur until sufficient housing was available for the displaced. Past recollections have shown them that all displaced persons were not adequately cared for. They needed some definite commitment beforehand.
3. Those in the lower income groups believed that regardless of their opinion, the government would do what it wanted to do.
4. Most fear a loss of neighborhood identification. This tended to be based on such selective features as the church, the block, neighbors of similar education and income, prestige or status of the area and on sentiment. It made them feel less secure.
5. A few hoped that whatever action taken would result in a better environment either in the existing neighborhood or a new one if relocation occurred.

There are many other social effects which result from highway improvements. Such improvements have resulted in a greater mobility of the society in general. Though such mobility may have some adverse effects by weakening family and community ties, it has increased the opportunity for more and more social interactions. Thus it has helped to reduce provincialism.

Highway improvements have facilitated the drift of population to metropolitan areas. Since 1950 about 90 percent of the growth of population in the U. S. has occurred in metropolitan areas.

Also by shortening travel time, highway improvements have led to the development of lands that were formerly too remote for residential use. These improvements have increased job opportunities and have resulted in savings to employees in time or money or both.

Highway improvements also affect the efficiency obtained from other public services. They make police officers more mobile by increasing coordination between state and local police and by permitting use of more centralized headquarters. In cases of emergencies (fire, medical aid, etc.) they have increased the chances of getting aid.

Social factors should be given some high priority in evaluating the impact of highway improvements. Highway improvements are made to satisfy human needs in one way or the other. Success in meeting these needs should, among other things, be evaluated against social values. A highway improvement causes some reordering or adjustment of the social structure. Unfortunately, however, there are no acceptable yardsticks for measuring these changes. As Professor Hennes pointed out, the lack of precise means for measuring changes does not excuse

highway engineers from exercising some intuitive judgement in the determination of social benefits or costs. To ignore realities because numbers cannot be found would be most unfortunate. (9)

C. The Political Impacts of Highway Improvements

The political implications of highway improvements are sometimes felt by the inhabitants of the affected area. A highway improvement might serve as a uniting force between several isolated political units by providing a link and reducing the friction of space between them. Such highway improvements also help to foster social interactions and understanding among people. They might enable people to have closer contact with their elected officials, increasing their participation in the political process.

While a highway improvement might in one case serve to bridge the gap between two or more political entities, it might under other conditions become a physical divider, disrupting community ties and causing changes in political boundaries. This disruption of boundaries could shift some taxable land from one area to another causing perhaps some significant changes in their economic bases. Thus, while it might boost the tax base of one given area, it could result in a net loss of tax money to another.

D. The Economic Impacts of Highway Improvements

Highway improvements are known to have a definite influence on the economic base of an area in different ways. By altering time - distance factors, they increase an area's potential to export goods and services. They facilitate access to markets and supplies, thereby

fostering the decentralization of industries. Highway improvements also tend to channel investment into nearby locations thus encouraging more basic economic activities in the area.

These economic impacts of highway improvements on their surrounding environment were well described by Richard Tybout.(10)

> "The history of economic growth testifies to the importance of improved transportation. At the earliest stages in which today's underdeveloped countries find themselves, there is little transportation. Each community is a relatively independent economic unit. As long as pack-horse transportation sets the cost of moving goods, there are very few goods sufficiently valuable to be carried overland. A second stage in economic progress occurs with the growth of improved and hence low-cost transportation. More goods can stand the cost of shipment, markets are broadened and more production can take place at the same location. Specialization and later mass production is made possible by trade which in turn is made possible by transportation.
>
> As long as the costs of transportation are covered in the final price of goods transported, a test of social desirability of transportation is readily available.----- Reorganization in the location and method of production occur as a result of improved transportation."

Thus, where favorable conditions exist highway improvements might provide the needed antidote to revitalize the economy of an area. In order to determine the benefits accruing from such highway improvements, three classes of benefits are usually identified: users' (or vehicular) benefits, nonusers' (or non-vehicular) benefits and reorganization benefits.

Vehicular benefits could be defined in terms of reductions in costs produced by highway improvements. These cost reductions are of four kinds: reductions in operating costs, reductions in time costs, reductions in accident costs, and reduction in the so-called impedance costs (i.e. the strains and discomforts of driving under congested conditions).

It is safe to say that these four classes of vehicular benefits do exist; the question is how to measure the savings brought about by highway improvements for each class. To attempt to answer this question one must be able to establish average values for each of these four costs for each size and weight group of vehicles under a variety of operating conditions that span the "before and after" states of highway improvement. Early studies conducted by McCullough, Beakey and their associates in the Oregon State Highway Departments (11) paved the way for deriving an acceptable technique for measuring these vehicular benefits. The Informational Report on Road Users Benefit Analyses for Highway Improvements (12) now provides a workable technique, though by no means perfect, of estimating vehicular benefits.

In addition to benefits enjoyed by those using a facility, there are advantages enjoyed by businesses and parcels of land because of their location near the facility. These advantages accruing to such economic entities as business, industries, the real estate market and the community in general are usually classed as non-user economic benefits.

Land value changes provide one of the indicators of non-user benefits. Land buyers expect to obtain certain present and future benefits from highway improvements. Interest in this aspect of economic impacts dates at least to 1927 when Robert M. Haig expressed the view that"

> "Transportation is in essence a method of overcoming the friction of space.... obviously an improvement in transportation, other things remaining the same, will mean a reduction in friction and

> consequently the diminution of the aggregate sum of site rentals. The two elements, transportation costs and site rentals, are thus seen to be complementary. Together they may be termed the costs of friction". (13)

A review of some economic impact studies which have been conducted since the time Haig made his statement will be made in a subsequent chapter of this work.

Reorganization benefits result from the reorganization or the relocation of activity following a highway improvement. Such a reorganization may be required in order to take advantage of a change in relative accessibility brought about by the improvement. These reorganizations may result in more profits for the industries concerned, or a lower unit price of products to the consumers.

It may never be possible to measure the aggregate impact of any highway improvement. But this limitation should not prevent the measuring of those aspects of the total impact for which some measures can be devised.

## Measurement of Economic Impact

Although most efforts at measuring the impact of a highway improvement have been directed at the economic aspect, there still exists a gap that must be filled before we can measure the total economic impact. Figure 3 shows that the victims and beneficiaries of highway improvement decisions can be classified either as users or non-users of the improvement. Most of the work done to date in the economic sector has dealt with the impacts of highway improvements on the users. Also in most decision-making, engineers, have been satisfied with computing the benefit-cost ratio, from the users' viewpoint. This approach is

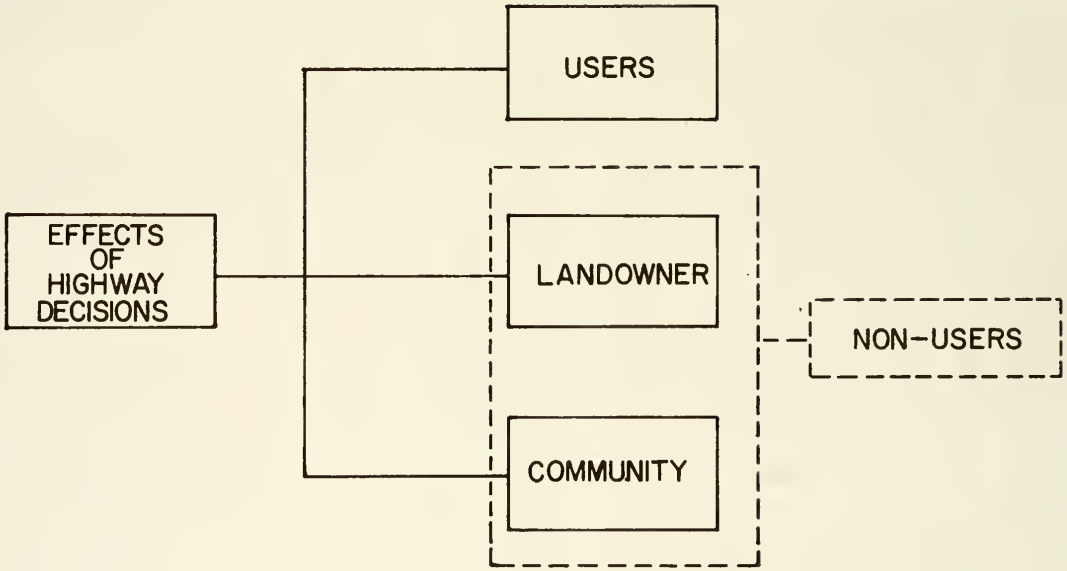



FIGURE 3: VICTIMS AND BENEFICIARIES OF HIGHWAY IMPROVEMENTS

inadequate, for no highway facility can be built in total isolation from its environment. The costs and benefits which such facilities transfer to adjacent land parcels and to the community in general must also be considered. Most previous attempts at considering the economic impact of highway improvements on land values were limited to descriptive before-and-after accounts of the observed impact. Such before-and-after measures of impact can of necessity be made only after the completion of the highway improvement. Unfortunately there are no generally accepted techniques for predicting the probable impacts of various highway decisions on land values at the time these decisions are made.

There is a dire need in many areas to fill this 'loophole'. Billions of dollars are being scheduled for investment in the development of our highway systems. An ever-increasing percentage of this sum will be spent on acquiring the necessary right-of-way. It has been estimated that of the total cost of the National System of Interstate amd Defense highways, eleven percent of the cost of the rural portions and 31% of the cost of the urban portions will need to be spent on right-of-way acquisition (14). Thus from a financial viewpoint it is only prudent to seek a means of inferring the impact of highway improvements on land values.

The need for such a predictive tool can hardly be overemphasized. Highway right-of-way agents need it to determine the influences of highways on land values. Real estate brokers and land developers need it to analyze highway impacts upon land development and investment prospects. City and county officials are concerned with the impact of

highway development upon property values which serve as the local tax base. As shown by the feedback loop in Figure 1, it is needed by the planning bodies who must consider the probable influence of highway development as a basis for future master planning.

Another significant way in which such a predictive technique would be useful is at highway hearings and in general public education. Richard M. Zettel, (15) expressed this strongly when he said:

> "...we often have to face up to the organized opposition, misguided or not, of businessmen and property owners whose economic concern in a particular highway undertaking is likely to be many times that of any individual highway user, even though the aggregate benefits to users far outweigh any possible economic losses to business and property. These directly-concerned business-men and property owners often raise a potent outcry against proposed highway improvements, sometimes with the sad results that the best highway plans are sacrificed for expediency."

Indeed we do need a quantitative and objective technique for predicting the impact of highway improvement on land values.

## Purpose and Scope

The purpose of this research was to alleviate an existing deficency in ascertaining total economic impact of highway improvements by developing a predictive model which would estimate the economic impact of highway improvements on land parcels. In this research, sale price is used as the indicator of impacts that occur to land parcels following such highway improvements. This decision was made because sale prices tend to be more objective than other indicators, such as opinion surveys. Sale prices avoid most of the problems associated with interpreting the real attitudes of respondents and they are based on verifiable contracts rather than estimates or responses that may some-

times be self-serving or otherwise inaccurate. (16)

The suggestion has long been made that changes in transportation improvements have effects on the value of land. Mr. A. Young, British Secretary of Agriculture, noted that:

> "The roadway was no sooner completed than rents rose from 7 shillings to 11 shillings per acre; nor is there a gentleman in the country who does not acknowledge and date the prosperity of the country to this road." (17)

Although this observation was made in 1813, little work has since been done, beyond mere descriptive accounts, to probe this relationship between the incidence of highway improvement and land value changes. This research, it is hoped, would at least spark some more work in this area with the ultimate goal of achieving an acceptable general predictive model for highway impacts on land value.

This study ulitized data from the states of Indiana and Florida. The Indiana data was obtained from the work done by V. G. Stover in his study of remainder parcels (18). The Florida data were made available to this study through the cooperation of the Economics Division of the Bureau of Public Roads. A possible difficulty with such data is that the original purposes of the studies for which these data were collected might have influenced the choice of source material and the accuracy with which they were collected, compiled and interpreted.

Based on these data, multiple linear regression predictive models were developed using the BIMD02R stepwise regression package program. Most of the independent variables in these models were qualitative in nature. These qualitative variables were introduced into the developed

multiple regression models through the use of dummy variables.

In addition to the development of these regression models, this research also investigated the analogy between the impact problem and the medical diagnosis problem. Techniques based on Baye's conditional probability were then applied to the solution of the impact problem.

# CHAPTER TWO

# A REVIEW OF THE STATE OF THE ART IN LAND ECONOMIC STUDIES

## General

The basic objective of economic research is to develop through investigation and careful study a body of fundamental knowledge essential to a proper understanding of the economic factors associated with the construction and improvement of highways. In this respect, land economic studies have helped to provide factual information on the economic effects accruing from highway improvements. The availability of such information for the use and guidance of highway planners, appraisers and right-of-way buyers increases the probability of the consideration of economic factors in highway location and design and provides the opportunity for more realistic and professional appraisals of properties acquired for highway purposes.

Land economic studies are more commonly regarded as embracing both economic impact studies and severance damage studies. These two types of studies are similar in that in both of them the researcher sets out to identify and measure the effects accruing from highway improvements. Both studies follow the "before and after" approach in estimating the impacts caused by the new improvement.

On the other hand some differences can be recognized between these two types of studies. While economic impact studies seek to identify the benefits or disadvantages affecting an entire community, severance studies concentrate on the effects of the highway improvement on particular land parcels taken in part for highway use. Therefore, in

severance damage studies, benefits to a given land parcel, that may not be shared by the community in general, are treated. Economic impact studies, on the other hand, are designed to identify and measure more general benefits.

Differences also exist in the approaches employed in these two types of studies. Highway severance damage studies often use a case study approach to identify changes occurring in the remainder parcel. Economic impact studies on the other hand are more concerned with land value changes and general economic changes in the affected area. The "control area" employed in economic impact studies is ideally one that is in all ways like the "study area" except that it is so located that it is not under the influence of the highway improvement. The criterion used in choosing a "control area" for severance damage depends to a large extent on the state laws regarding compensation. "In states where both general and special benefits can be applied against the cost of acquiring right-of-way property, a control area removed from the highway influence is desirable. However, in more than half of the states where only special benefits are to be considered in the determination of adjustments to be made with affected property owners, control areas are needed in the immediate neighborhood of the study parcel." (1)

## A Review of Highway Severance Damage Studies

The term "severance damage" implies that by severing a portion of land from its parent tract, the parent tract suffers a loss in value. Evidence from many severance damage studies in the United States, however, tends to discount this assertion. They show to the contrary

that most remainder parcels are in fact enhanced in value. Thus, the implication of the above term can be regarded as unacceptable.

In exercising the power of eminent domain, the partial taking of a property usually presents the most challenging problem of arriving at just compensation for the part taken plus damages to the parcel remaining. Severance damage studies are intended to provide to those concerned the basic data needed to arrive at an equitable payment for the property taken.

Generally two causes of severance damages resulting from the taking a portion of a tract of land for public use can be recognized. First, a severance damage can occur as a result of a change in the use of the residual property to less profitable use as evidenced by a decreased rental and consequent decreased market value. This condition may be brought about by:

a. A resulting insufficiency of the remainder to support the normal enterprise,

b. A resulting distorted shape,

c. A resulting loss of railroad trackage from a portion of the remaining parcel in industrial property,

d. A change in the grade of the street and

e. Any condition causing an injurious physical interference with the use of the remainder ( 2 ).

A second cause of severance damages is the burden (with respect to cost) imposed upon the residual property occasioned by the taking of a portion of it and the construction of the improvement in the manner proposed such as:

a. The increased cost of construction of an irrigation system, if and when built

b. The cost of fencing the residual land along the land taken, if necessary

c. The increased cost of construction of a building because of the unusual shape or special foundations required.

In short, the damage contemplates a physical interference with the most profitable use of the property and excludes such considerations as fear of unsightliness. To assert just compensations for such damages, many states have completed or are completing severance damage studies. Figure 4 shows that up to 1963, 46 states in the United States had completed or were undertaking severance damage studies. Results from these studies covering over 1,200 cases were submitted to a U. S. Bureau of Public Roads central file. In addition to these case studies, the states had, as of 1963, issued more than 1,500 individual case study reports.

The need for such a central bank for severance case data was stated by Rudolf Hess of the California Division of Highways before the American Institute of Real Estate Appraisers meeting in Los Angeles in November, 1964, in these words:

> "The remainder parcel sales data collection is one of the first direct research links between the appraisal profession and land economics. The analysis of remainder parcel sales data may provide the vehicles for a firmer marriage of these two specialized areas; it may provide the means for appraisers to more conveniently and confidently consider land economic data and theory. Collective analysis is the key. Collective analysis is a tool of land economics---------."

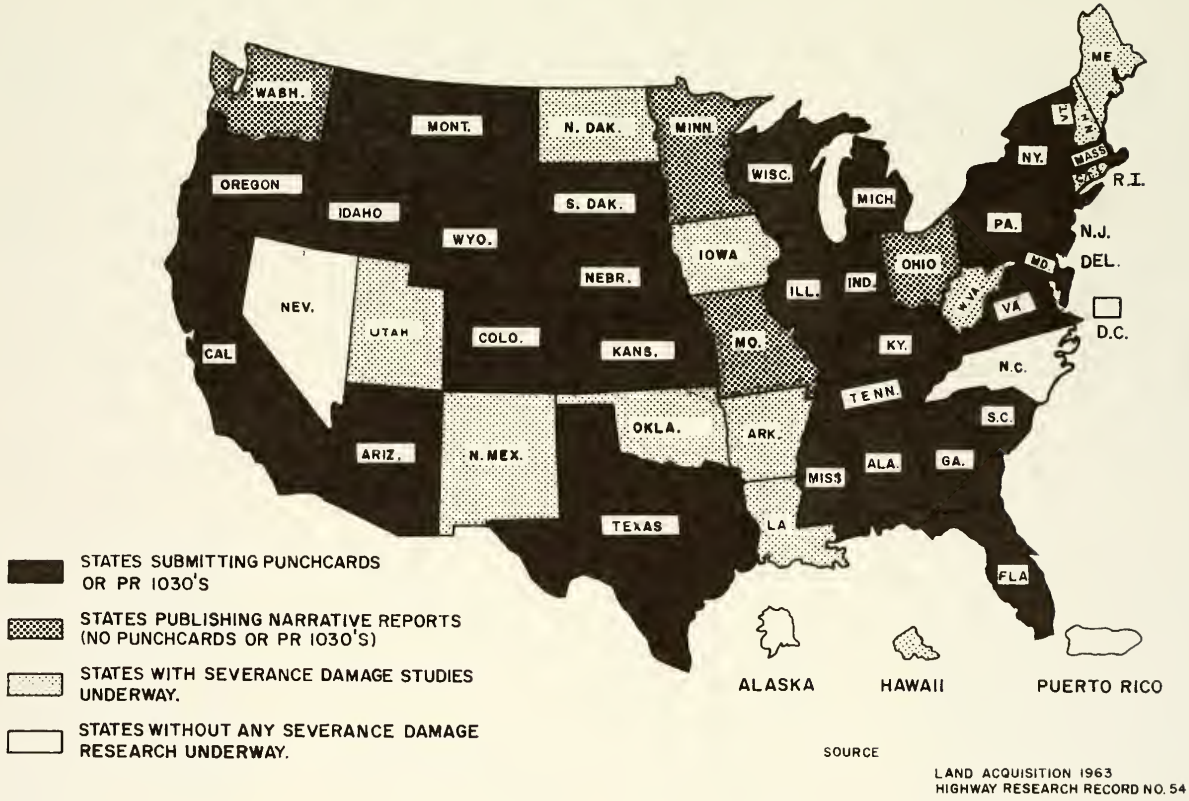



FIGURE 4: STATES CONDUCTING SEVERANCE DAMAGE STUDIES.

However, prior to his statement the Bureau of Public Roads had in 1961 started to collect information on the effect of taking a part of a property for highway right-of-way in a Severance Effects Bank. By April, 1967 the Public Roads Severance Bank had collected about 3,000 cases covering 4,000 sales of remainders from these cases. Also about 2,900 narrative cases were available. The use of computers in recording cases has facilitated the handling of this vast amount of data and has provided easy access to cases for future use.

## Methodology in Severance Damage Studies

In determining the extent to which the owner is "made whole", the before value of an entire tract is compared with the total amount received by the owner (from payment for property taken plus any payment for damages plus the sales price of the entire remainder).

To determine the after value (AV), the value of the property at time of taking (BV) as indicated by staff appraisers, fee appraiser or both is used. From this figure is deducted the just compensation (JC) paid to the landowner for property and damages as indicated by the settlement. This formula, with all terms in constant dollars, is as follows:

$$AV = BV - JC$$

The total land value benefit ($B_T$) or damage ($D_T$) is determined by using the subsequent sale price (SP) as indicative of the market value, and from this amount is deducted the after value (AV) if less than the sale price. If more than the sale price, the sale price is deducted from the after value to show the damage. In other words,

two cases result:

Case 1: $SP > AV$

Then benefit, $B_T = SP - AV$

Case 11: $SP < AV$

Then damage, $D_T = AV - SP$

If $B_H$ represents the benefit accruing from the highway improvement and $B_G$ a general increase in value for all parcels not in any way due to the improvement, we have:

$$B_H = B_T - B_G$$

When local indices of increase in real estate market are available, these should be used to estimate the magnitude of $B_G$. When not available, national indices may be used.

## Findings from Severance Damage Studies

An analysis of the data ( 3 ) collected in the Public Roads Severance Effects Bank reveals that:

1. In most cases, the selling price (SP) of land parcels exceeded the after value (AV). Figure 5 shows that approximately 75% of all cases in the bank had a recovery rate greater than 100%, in other words, selling price was greater than after value.
2. The time of sale was found to have some effects on the recovery rate. This effect of time is of interest because it affects comparisons of before and after values. Figure 6 shows how recovery rates changed with time of sale for the cases in the Public Roads bank. General increases in land value were not isolated in these cases. Those selling a year

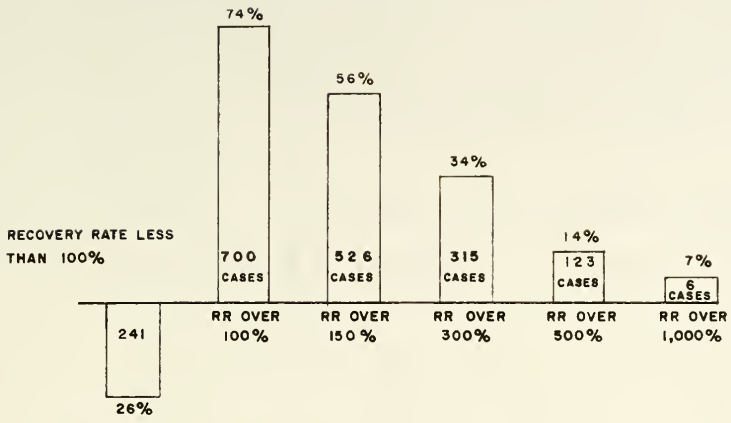



FIGURE 5: LAND VALUE RECOVERY RATES (OVER ALL), BY NUMBER AND PERCENT OF CASES.

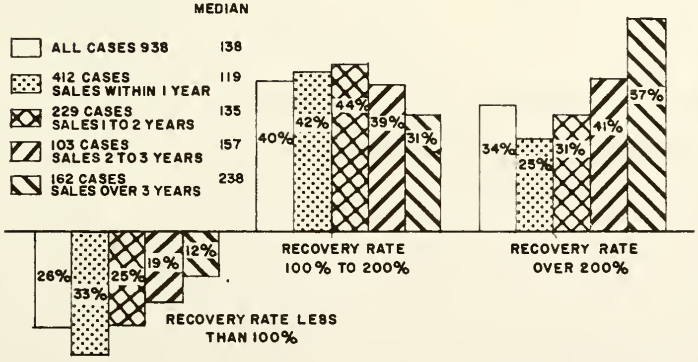



FIGURE 6: LAND VALUE RECOVERY RATES, BY TIME FROM ACQUISITION TO SALE, UNADJUSTED FOR GENERAL LAND VALUE CHANGES.

or more after had a higher recovery rate. This emphasizes that when sales of remainders are made after a lapse of time (a year or more) after the taking, adjustment should be made for general land value changes which are not due to the highway improvement.

3. Land use of the property at the time of taking appeared to influence the recovery rate. This is shown in Figure 7 . Owners of vacant property are more likely to gain more from the sale of remainders than owners of residential or agricultural properties. Information as of 1963 on percentage of owners paid damages and the percentage having apparent loss categorized by land use before the highway taking was:

| | |
|---|---|
| Vacant: | 49% and 12% |
| Agricultural: | 67% and 26% |
| Residential: | 67% and 22% |

4. Some significant variations existed between the recovery rates of different types of remainders. The three main types of remainders were thus defined:

   a. A separated parcel is the remainder containing the improvements. A separated parcel results when a highway taking leaves two remainders or when only one parcel remains - a situation often referred to as "Severed".

   b. An isolated parcel is an unimproved remainder which generally can be reached only by a public road.

   c. A parcel is landlocked when no access to the parcel

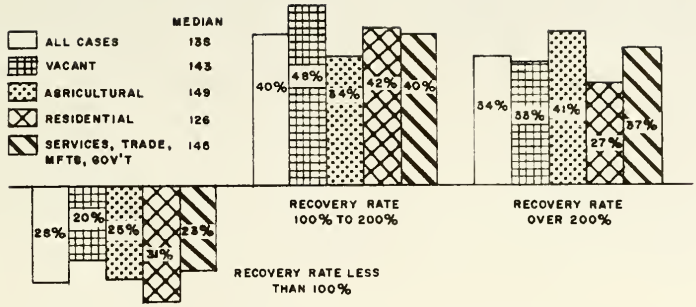



FIGURE 7. LAND VALUE RECOVERY RATES, BY LAND USE AT TIME OF ACQUISITION.

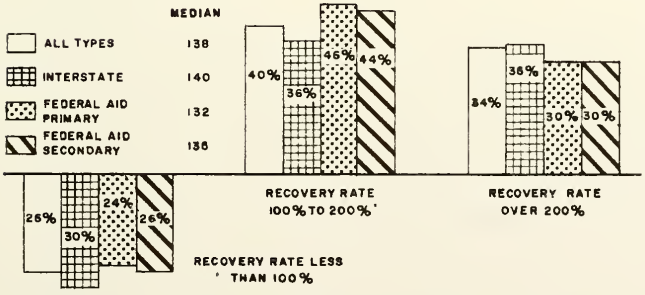



FIGURE 8. LAND VALUE RECOVERY RATES, BY TYPE OF HIGHWAY SYSTEM.

exists by use of public facilities or adjacent land of the same owner.

Results show that 18% of the separated parcels, 35% of the isolated parcels and 54% of the landlocked parcels had a recovery rate less than 100%. On the other hand, 38% of the separated parcels, 24% of the isolated parcels and 14% of the landlocked parcels had recovery rates greater than 200%. These percentages were based on the few cases available in the bank as of 1963.

5. The type of highway system adjacent to the remainder parcel appeared to influence the recovery rate slightly. The median recovery date for remainder parcels along interstate routes for data in the bank as of 1963 was about 140%, 2% higher than the median recovery rate for all cases in the Bureau's Bank. At the same time, recovery rates of 132% and 135% were calculated for cases along Federal-aid primary roads and Federal-aid secondary roads respectively. Figure 8 shows the above variation of land value recovery rates, by type of highway system.

6. Analysis of recovery rates by visibility showed that parcels fully visible from the highway had a median recovery rate of 145% compared with recovery rates of 133% for parcels from which the highway was partially visible and 117% for parcels from which the highway could not be seen.

7. The travel distance from the remainder to the new highway appeared to have some bearing on the recovery rate of the

remainder parcel. Figure 9 shows that 37% of the remainder parcels within 1/2 mile travel distance of the highway had a recovery rate greater than 200% while only about 25% of those parcels with longer travel distances had such a recovery rate. Only 21% of the remainder parcels within 1/2 mile of the new highway had recovery rates of less than 100%; for remainder parcels more than 1/2 mile in travel distance from the highway, about 42% had recovery rates of less than 100%.

8. The proximity of a remainder parcel to an interchange appeared to have a significant bearing on its recovery rate. The median recovery rate for parcels located close to an interchange was found to be about 164% compared with 131% for parcels located away from the interchange (Figure 10 ).

9. For cases in the Bureau Bank as of 1963, the median recovery rate was found to be 138%. This appears high. It should be remembered, however, that recovery rates are on a per acre basis and that in many cases only a portion of the remainder is sold. Also on the whole, analysis showed that the owner was being made whole in only four out of five cases.

10. Studying the simultaneous effects of all above factors on the recovery rate through multiple regression techniques revealed that the most influential factors were:

    a. Change in land use

    b. Time elapsing from acquisition to sale

    c. Travel distance to new highway

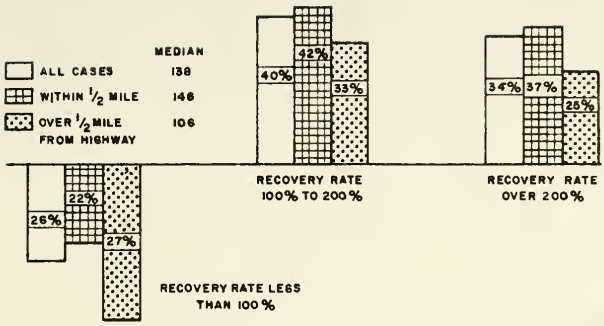



FIGURE 9: LAND VALUE RECOVERY RATES, BY TRAVEL DISTANCE TO NEW HIGHWAY.

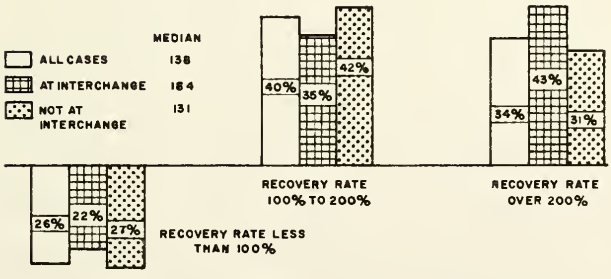



FIGURE 10: LAND VALUE RECOVERY RATES, BY NEARNESS TO INTERCHANGE (OVER-ALL).

d. Type of remainder

e. Nearness to interchange.

A study of remainder parcels conducted by Mr. Vergil G. Stover ( 4 ) in the State of Indiana showed that there was in general an overpayment for damages for highway right-of-way in the State of Indiana in the 1955-1961 period. Remainder parcels on secondary and primary routes were enhanced to a greater extent than those parcels along interstate routes. A high percentage of the tracts which continued on residential use after the taking were damaged and relatively few were enhanced. The author observed that in most cases where the residual was enhanced, the grantee anticipated utilizing the tract for some purpose other than the existing use. His work also showed that more damages were paid for landlocking and for separation of properties than were sustained. This is in agreement with observations at the national level.

B. A Review of Economic Impact Studies

The Highway Revenue Act of 1956 gave a big boost to economic impact studies of highway improvements. This boost is revealed by the fact that most of the economic impact studies completed in the United States during the last twenty years were conducted after the passage of this Act. Most of these studies sought to achieve one of the following objectives; to collect primary data from a given location in order to determine the economic effect of a particular improvement, to determine the impact of a highway improvement on industry or some other sector of the economy or to investigate theoretically the consequences following highway improvements. Other studies measured the

impact in the area being bypassed by a new highway. All these studies have helped to provide a representative view of highway influences on the community in general.

Benefits accruing from a highway improvement may be categorized into three classes: User benefits, non-user benefits and reorganization benefits ( 5 ). User or vehicular benefits represent those that are enjoyed by the highway users only. These are cost reductions enjoyed by the highway users as a result of the improvement. Non-user benefits are those that are enjoyed by other indirect beneficiaries of highway improvements. These include advantages to land use patterns and to land values. Reorganization benefits are defined as those accruing to a community from the reorganization of activities following the construction of a highway improvement.

The remaining discussion here will be devoted to a description of the economic impact of highway improvements on land value, a non-user benefit.

## Techniques in Economic Impact Studies

Considerable differences exist among the techniques employed in various studies to filter out the influence of a highway improvement on land values. The "before and after" approach is the most popular technique. This technique attempts to obtain an indication of the changes in land values by comparing property values before the new highway improvement was built with property values after the construction of the improvement.

Some studies attempted to extract the full implications of land value changes by using a technique involving "study" and "control"

areas. The "control area" here has the same connotations as in biology and some social sciences. It denotes a group that can be assumed to be untouched by whatever treatment the "test group" is subjected. To the degree that these studies were able to choose an appropriate "control" area, the measured land value differentials provided an indication of the portion of the change that was attributable to the new highway improvement.

An analysis conducted by the Secretary of Commerce in 1961 revealed that 56 economic impact studies of urban land values conducted in 19 states were comparable enough to be analyzed. These 56 studies covered 183 study segments of which 151 were classified according to the type of land use. Fifty-one of the study segments used, both "study" and "control" areas so that differences in percentage changes could be observed. Of the 51 study segments reporting changes in both study and control areas, fifteen showed the changes in land values separately from changes in the values of the building.

## Findings of Major Economic Impact Studies

To provide a comparable basis for the various findings of major economic impact studies in the United States, the Secretary of Commerce ( 6 ) presented all land value changes in constant dollars and on an annual rate of change. An analysis of these findings led to the following conclusions:

1. Only 11 of the 183 study segments showed a decrease in land values and only one study segment reported a decrease of more than 5% per annum.
2. The rates of increase of land value varied considerably according to land use. Industrial land values were found

to have increased the most. The land use classification in all cases was dependent on the use at the end of the study period and much of the land classified as industrial land had been devoted to agricultural purposes prior to construction of the new highway and had shifted to an industrial use only because of the new highway.

3. Thirty-three of the study segments reported in House Document 72 were commercial land uses. These study segments showed that the average land price increase was between 10% and 11% per year.
4. The largest number (85) of the study segments were pertinent to determining the changes in residential land prices. The average annual increase in this category was between 8% and 9%.
5. The 22 study segments reporting changes in land value of unimproved land showed an average increase of between 12% and 13% per year. Two of these segments showed annual increases of more than 65%.
6. The median annual rate of increase in land prices for the 183 individual study segments reported in House Document 72 ranged from 6% to almost 18% which was 2 to 6 times the annual growth in gross national product as of 1961.
7. Land values of parcels closest to the improvement showed the greatest increase. Land value changes along the Gulf Freeway provide a good illustration of this point. The Gulf Freeway Study revealed that land values in two zones near the Freeway reached 667% and 242% of what the average values had been originally, whereas land parcels in two control areas were

valued at 80% and 203% respectively of their former values.

In most of the economic impact studies conducted in the State of Indiana (7,8 ) the "band concept" was used in conjunction with the "before and after" technique. The "band concept" permitted the analysis of land value changes at various distances from the highway improvement.

Studying the impact of the Kokomo bypass from 1950 to 1964, Mr. E. G. Evans observed that land values adjacent to the bypass increased at a faster rate than land values at other locations in the Kokomo area. Moreover, for several cases he observed that the land values of properties adjacent to the bypass increased significantly with time. The analysis of land value data by R. J. Hensen ( 7 ) and others in the Lebanon bypass study also revealed that land parcels close to the bypass experienced a rapid and considerable increase in value. The results of these studies are discussed in Chapter Three.

## Impact Research at Purdue University

The finding of three research studies in the severance damage and highway impact areas which were conducted at Purdue University have already been referenced and discussed. These studies constitute a part of the impact research at Purdue University, of which this present study is also a part. In the next several paragraphs the history and scope of the total impact research at Purdue University will be reviewed.

The Advisory Board of the Joint Highway Research project at Purdue University on August 17, 1950, recommended in a report to the Indiana State Highway Commission that a study be undertaken to determine the

effects of a bypass upon a community. The commission, in accepting this recommendation, authorized studies of the Kokomo and Lebanon bypasses to be made jointly by the Metropolitan Area Traffic Survey unit of the Commission and the Joint Highway Research Project of Purdue University ( 9 ). The initial studies were performed during the period September 1950 to July 1951. Since both by-passes were under construction and completed during this period, studies included just before and immediately after conditions. The location of these bypasses with respect to the State of Indiana is shown on Figure 11.

On July 1, 1960, a ten year program on impact research was undertaken by the Joint Highway Research Project at Purdue University. This project was designed to provide information on the effects of highway improvements on adjacent areas, including land use changes, land value changes and changes in the characteristics of highway travel. Such information would provide for a more efficient and economical approach for locating future highways.

A total of six specific types of highway improvements were proposed to comprise the study areas. These facilities all of which are major state Highways are as follows:

Facility 1: an urban by-pass with complete access control.

Facility 2: a rural highway with complete access control.

Facility 3: an urban by-pass with little or no access control.

Facility 4: a rural highway with little or no control.

Facility 5: a bridge and its approaches in an urban area.

Facility 6: a major highway interchange near a metropolitan area.

Facilities chosen to correspond to the types of improvements listed

above are (See Figure 11 ):

Facility 1: The Interstate 65 bypass around Lebanon, Indiana;

Facility 2: A thirteen mile portion of Interstate 65 from the south end of the Lebanon bypass to the interchange with Interstate 465 northwest of Indianapolis, Indiana.

Facility 3: The U. S. 31 bypass around Kokomo, Indiana.

Facility 4: U. S. 31 from the south end of the Kokomo bypass to the north edge of Marion County, Indiana.

Facility 5: The U. S. 231 Bridge over the Wabash River connecting Lafayette and West Lafayette, Indiana.

Facility 6: The interchange connecting Interstate 65 and Interstate 465 northwest of Indianapolis, Indiana.

In October 1964 a report was submitted by R. J. Hensen and others on Facility 1 ( 7 ). A study of facility 2 was conducted and submitted in June 1961 by Fletcher ( 10). A report on facility 5 was completed in May 1962 ( 11). In August 1965 a report on facility 3 was submitted ( 8 ). Fleischman ( 12) submitted in June 1968 a report on a further study on facility 5.

Before the initiation of this 10 year program H. L. Michael in 1951 ( 9 ), A. D. May in 1957 ( 13 ), Charles Pinnel in 1958 ( 14 ) had completed three reports on the effects of the Kokomo and Lebanon bypasses.

In addition to the above list of impact studies V. G. Stover conducted research on remainder parcels as part of this total impact research effect in the State of Indiana ( 4 ). The intent of this research was to supply information on the effects of Indiana's highway

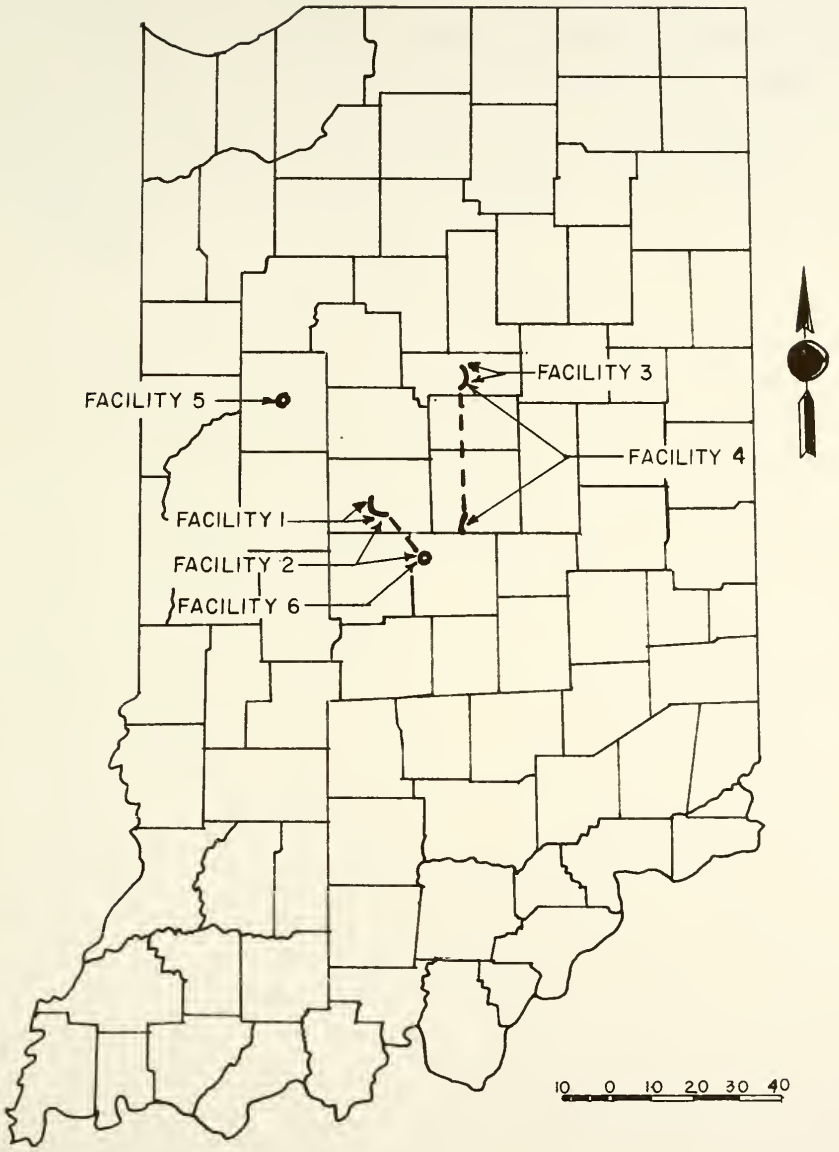



FIGURE II: LOCATION OF STUDY FACILITIES IN INDIANA

improvements, especially data that would be useful to the divisions of planning and of land acquisition of the Indiana State Highway Commission.

The research presented in this current report culminates the work on the economic impact initiated in 1960. Feeding on the results of the previous studies made at Purdue University and elsewhere, an attempt will be made at modeling the impact of highway improvements on land value.

# CHAPTER THREE

## AN 'ANATOMY' OF LAND VALUE CHANGES

The land market is a complex one. Unlike other markets where one knows the product and the supply and demand relationships, the forces operating in the land market are not so clearly defined. The complexity in the land market arises among other things from the problem of determining the amount of the land and building return that constitutes the total real estate value of a property. An understanding of the real estate market becomes even more difficult because essentially the value of a piece of real estate lies in its future productivity of service. Knowing the productivity of service at a future date entails the art of prediction. In economics, the value of any capital good is the discounted present value of the future returns. Therefore to do a good job at estimating land values, one needs to more fully understand the factors influencing land value fluctuations from one time period to another.

A glance at the population distribution of the United States suggests that physiographic conditions do influence the demand for land. People in general tend to congregate in large numbers in areas where physiographic conditions are conducive to living. This attractiveness of some areas relative to others for residential, industrial and other purposes constitutes a force which can sway the value of land. The operative zoning policies in an area could also lead to variances in the land values at different locations. Since demand and

supply forces act simultaneously to produce the price of a given piece of land, it is conceivable that a significant population shift in an area would affect the value of real estate in the area.

Apart from such population shifts the value of a real estate property is also affected when the price level rises or falls quite apart from the current level of supply and demand for land and buildings. Considering for a moment that physically a building is merely an accumulation of materials, then a change in the cost of these materials could affect the cost of the building. Therefore, if the prices of steel, sand, gravel, cement and lumber decline, then their dollar values also decline whether they are piled in a building supply dealer's yard or assembled in the form of a building. Similarly a change in land value can occur because of a change in the values for the uses for which the land is used.

A highway also can have an effect on land value. The comparative accessibility values of sites within its zones of influence might be greatly increased, thus channelling the demand for land along its length. This type of "gravitational" pull sets up a chain reaction: the highway improvements attract more factories into the area, the factories attract more employees who increase the demand for homesites. With this increase in demand for land, the forces of demand and supply operate on the land market to produce an increase in land value.

In impact research we seek to isolate that portion of the observed change that can be attributed to the highway improvement. Thus, in the "before and after" technique, adjustments must be made for general fluctuations in the economy as a whole as well as for changes in land value resulting from non-highway causes. Such adjustments are effected

by using study and control areas and examining the surrounding area to establish local trends of real estate values for lands similar to the ones involved in a given situation. An increase of 300% in the value of a property from the "before" period to the "after" period is not significant to our analysis of impacts due to highway improvements, if there were a general increase of such a magnitude in all real estate values. On the other hand, if local land values were rising at only 5% per year and 15% increase pear year was registered for the affected area, then certainly this difference could be attributed to the highway improvement. Sometimes however, local indices might not be available to appraise the local trend in the real estate market. In such cases, the general trend in the national economy as indicated by the consumer price index or other indices might provide a clue to that portion of land value change caused by inflation. Figure 12 shows the general price trends for the years 1945 to 1967. The consumer price index permits a reasonable adjustment for price fluctuations due to inflation, thus affording a comparison on a constant dollar basis.

The rest of this section will be devoted to an examination of the factors that appear to be related to observed changes in land value that follows a highway improvement.

## Land-Use

Changes in land use and land value, as shown in the previous chapter, appear to be highly related. One exerts an influence on the other. However, the relationship between land use and land value is a curious one. Although a prospective change in one will generate a change in the other, land use and land value are not wholly

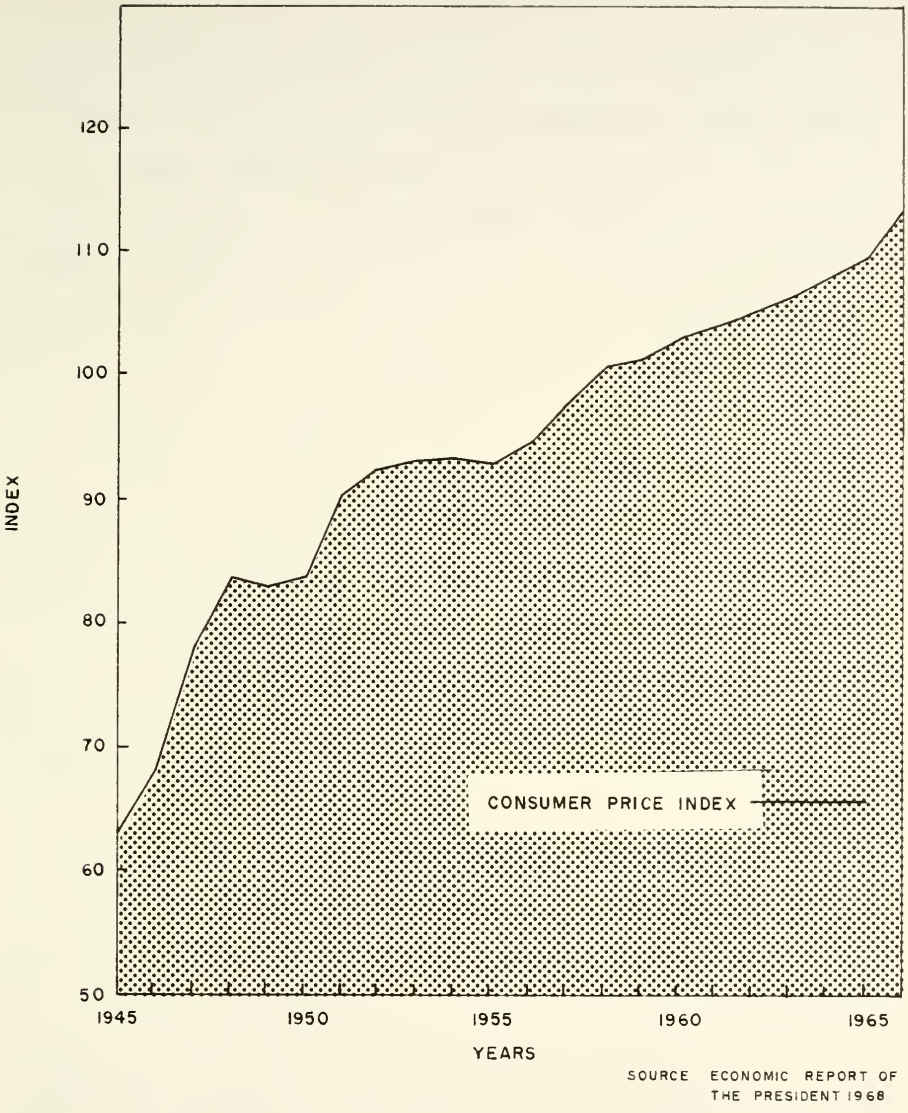



FIGURE 12 GENERAL PRICE TRENDS

interdependent. Shifts in land values may occur even when the prospective use of the land remains fixed. Also changes in land uses come about more slowly when economic conditions change than do changes in land value. This is accounted for by factors such as zoning regulations, availability of building capital, building permits, individual prejudices and reluctance to change. Also nationwide economic conditions may foster or deter such immediate changes. These inhibiting factors will vary from one community to another and even more among localities within a community.

The observed differences of land value changes for various land uses could be due to the fact that factors influencing land value have different significances for different land use types. For instance, accessibility may be important to the farmer for one reason, to the owner of a supermarket for another reason and to the suburban commuter for still another. Thus, the weight given to accessibility in these different decision situations varies.

## People's Attitude

Land value depends in an important way on how people feel about property and the surrounding area. Without public acceptance, a highway improvement might generate reductions in land values. The prospective land buyer considers many factors, consciously or unconsciously, in his mind before deciding on how much he would like to pay for a given piece of land. The buyer therefore must compromise between such factors as noise and dust, inconvenience of being adjacent to a freeway, future benefits to be derived from such a location, etc.

Resident opinion surveys carried out in New York (Figure 13 ), Texas ( 1 ) and California ( 2 ) provide a general clue to the attitudes of people to highway improvements.

It is not, however, possible to relate the intensity of people's feelings about these highway improvements to the observed land values in any precise way. It appears, nevertheless, that people's opinions about highway improvements play a role in determining land values.

## Accessibility

A distinction must be recognized between the terms 'access' and 'accessibility.' Access refers to a particular means of ingress and egress between a highway and abutting land, and accessibility refers to the relative proximity in travel time of land being evaluated to major reference points. Market data from sales and resales of residential sites abutting controlled-access highways strongly suggest the recognition of accessibility rather than direct access as the more important upon value for residential land use ( 3 ). In fact, some types of land use do not require any direct access and can be utilized with advantage by circuitous access to the adjacent throughfare via secondary streets and roads.

In a Washington Study ( 4 ) conducted by R. G. Hennes and others, the effects of accessibility on land values were investigated. As a result of the construction of the Lake Washington floating bridge connecting Seattle and Mercer Island, the travel time between these two ends was reduced from 45 to 15 minutes. To analyze the resulting impact, the sales prices of properties for pre and post bridge periods for the study area (bridge affected) and a control area were collected.

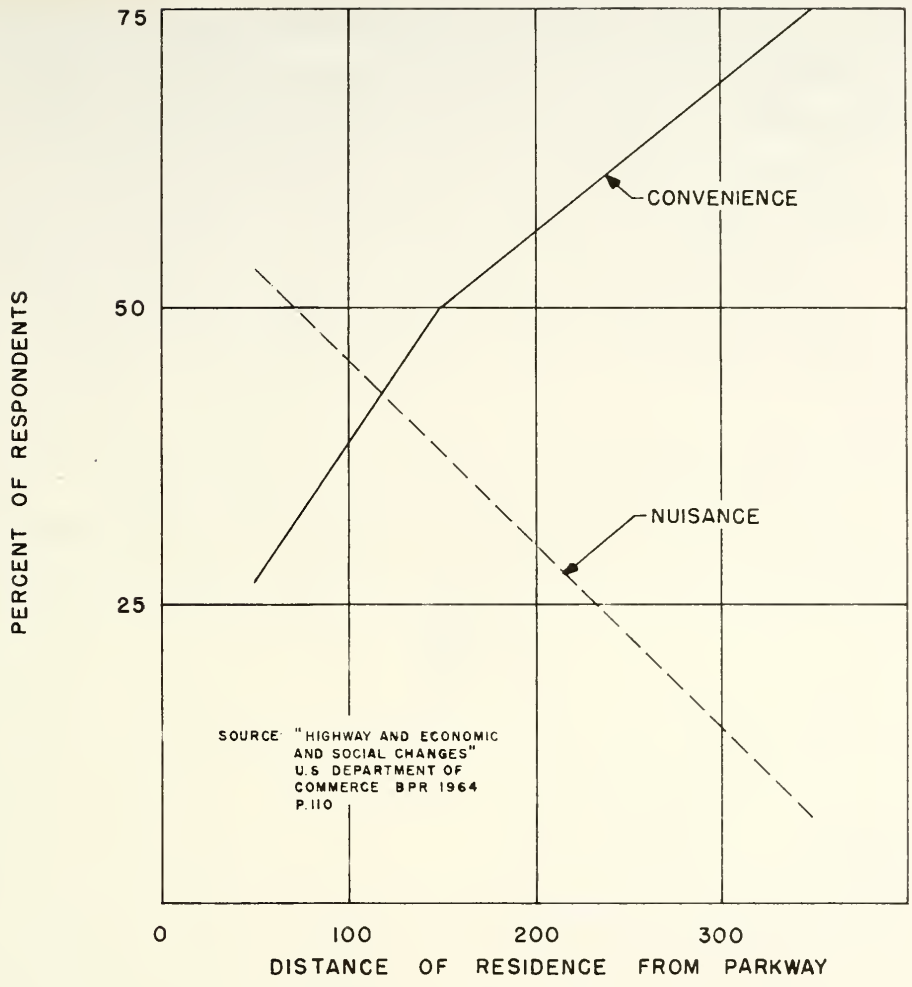



FIGURE 13: OPINIONS OF RESIDENTS OF NEARBY HIGHWAY IN WESTCHESTER COUNTY, NEW YORK (TRAFFIC IMPACT)

The obtained results revealed that this improved accessibility generated increases in residential land values of about 155% and 78% in areas of Mercer Island and in an area east of the lake, respectively.

This and other studies have shown that accessibility is one of the chief determinants of residential location. The results of these studies suggest that changes in land value can be said to reflect differences in accessibility and that it may be possible to use this information in estimating the effect on land value of highway improvements.

A simple model based on the miltiple reference point concept has been suggested by Garrison ( 5 ) to examine the value of accessibility to parcels of farm and nonfarm land. This model is of the form:

$$Y = \sum_{i=1}^{n} B_i X_i$$

where $B_i$ the weight attached to each unit of distance.

$X_i$ the distances to the n reference points of the residential site.

$Y$ the value of the parcel.

Using statistical procedure, the researcher can then determine how much of the observed variation in the values of the land parcels can be explained in terms of their different locations. This technique can also reveal how much each unit of distance, in terms of the $X_n$ reference points, contributes to Y. Applying this simple model to relate land values to accessibility (measured by the airline distance from 8

reference points) in the city of Spokane, Washington, revealed that simple linear distance measurements did not provide a satisfactory explanation of the observed variation in land values.

The use of airline distance, however, may be unrealistic as it does not represent the way people actually move about. Measuring accessibility in terms of the distances along the road network of a study area would be more representative and hopefully produce better correlation. Even then it seems more logical to use the travel time between different points in the network rather than distance in measuring this accessibility. Thus if there are six selected reference points i, j, k, l, m and n, travel time from a given land parcel to each of these reference points before and after the highway improvement can be determined. An accessibility index for parcel p at time t would then be:

$$A_t^p = \sum_q^6 T_{pq}$$

where $T_{pq}$ = travel time between parcel p and reference point q.

Then the change in accessibility potential of a parcel brought about by a highway improvement may be given by:

$$A_{\Delta t}^p = A_b^p - A_a^p$$

where $A_b^p$: accessibility index of parcel p <u>before</u> highway improvement

$A_a^p$: accessibility index parcel p
after highway improvement.

This change in accessibility index might serve as a useful indicator of changes in land value (especially for residential land uses) following a highway improvement.

The Distance of A Land Parcel From The Highway Improvement

The "Band Concept", employed in the impact studies conducted by Purdue University researchers for the State of Indiana, provided a means of investigating the relationship between land value and the proximity to the highway improvement.

Employing this concept, R. J. Hensen and others ( 6 ) investigated the impact of the Lebanon Bypass on the surrounding area. In this study four bands (A, B, C, D) were defined as follows:

| Band | Distance from facility in miles |
|---|---|
| A | 0 - 1/2 |
| B | 1/2 - 1 |
| C | 1 - 1 1/2 |
| D | 1 1/2 - 2 |

The unit value of land within any given band was assumed equal to the mean price per acre paid for land within that band. Sale prices were inferred from the Federal Revenue Stamps evaluation recorded on each deed. Using the sale price data from this study and adjusting it with the consumer price index with 1960 as the base year, the mean prices shown in Table 1 were obtained.

The results of this before and after analysis revealed (Figure 14 ) that changes in land values decreased as distances from the highway

TABLE 1 :MEAN PRICE PER ACRE ALONG LEBANON BYPASS IN CONSTANT DOLLARS

| BAND | BEFORE (1945-50) | AFTER (1951-56) |
|---|---|---|
| A | $370 | $1030 |
| B | 288 | 456 |
| C | 230 | 258 |
| D | 220 | 262 |

TABLE 2 :MEAN VALUE PER ACRE OF LAND ALONG KOKOMO BYPASS IN CONSTANT DOLLARS

| BAND | BEFORE (1945-50) | AFTER (1957-63) | % CHANGE |
|---|---|---|---|
| 1 | $349 | $1540 | 330 |
| 2 | 318 | 1850 | 480 |
| 3 | 549 | 2960 | 430 |
| 4 | 263 | 1400 | 440 |
| 5 | 115 | 415 | 260 |

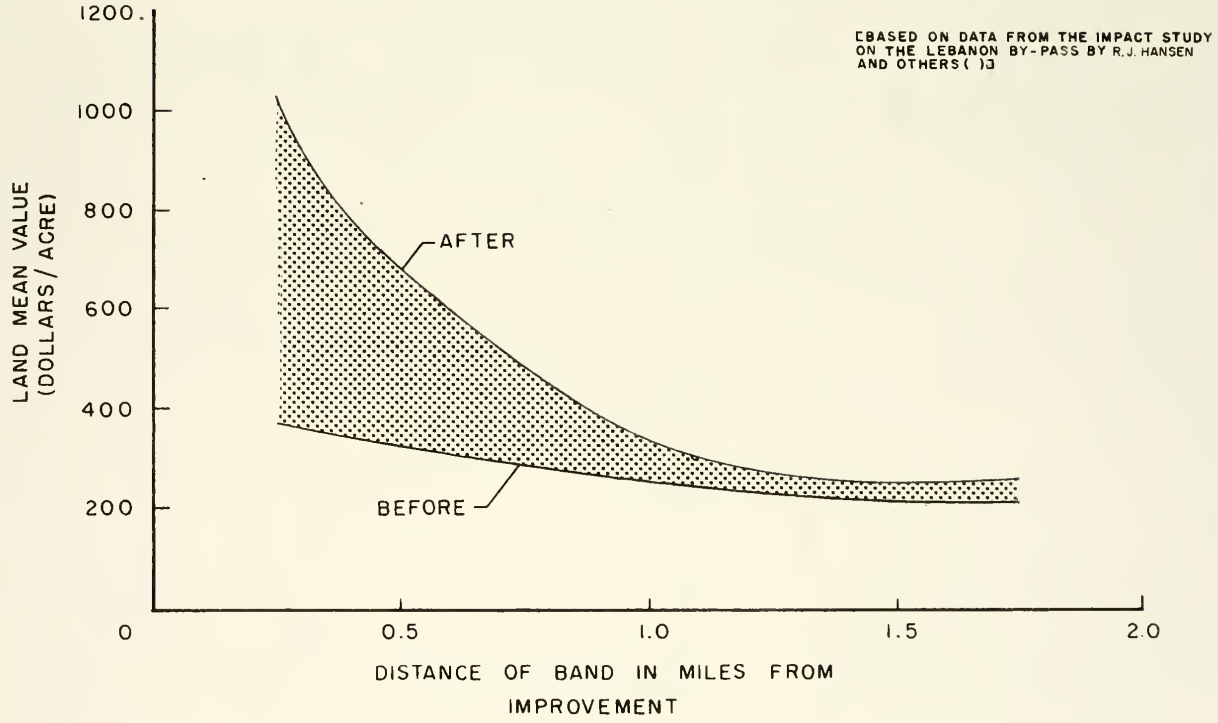



FIGURE 14: MEAN LAND VALUE (DOLLARS/ACRE) IN CONSTANT DOLLARS AT VARIOUS DISTANCES FROM THE IMPROVEMENT.

improvement increased. In other words, land parcels closest to the improvement experienced the greatest change in value.

This above approach was also used to investigate the effects of the Kokomo Bypass. Table 2 and Figure 15 show the mean value per acre of land along the Kokomo Bypass. In this case changes in land value on each side of the improvement were reported separately. The width and location of these bands are shown in Figure 16 .

Other impact studies in the United States confirm the basis observations made in these Indiana Studies. Land value changes along the Gulf Freeway provide a further illustration of this point and were presented in Chapter two.

The changes on both sides of an improvement may exhibit a "bell-shaped" relationship with symmentry about the zero distance line. However, this symmentry will not usually exist in practice. This is because two similar properties equidistant from a highway improvement but on opposing sides do not necessarily experience the same extent of change in land values. Other variables such as accessibility, distance from city center, type of land use, intensity of land use and levels of existing economic activities on both sides of the improvement contribute to the observed change in land value. Results from the Kokomo Bypass study shown in Figure 15 is an example of a case where symmentry was not attained around the zero line.

Distance From Nearest Urban Area

The distance from an urban place is probably an important factor in determining what will happen to a community or a land parcel subject to the influence of a highway improvement.

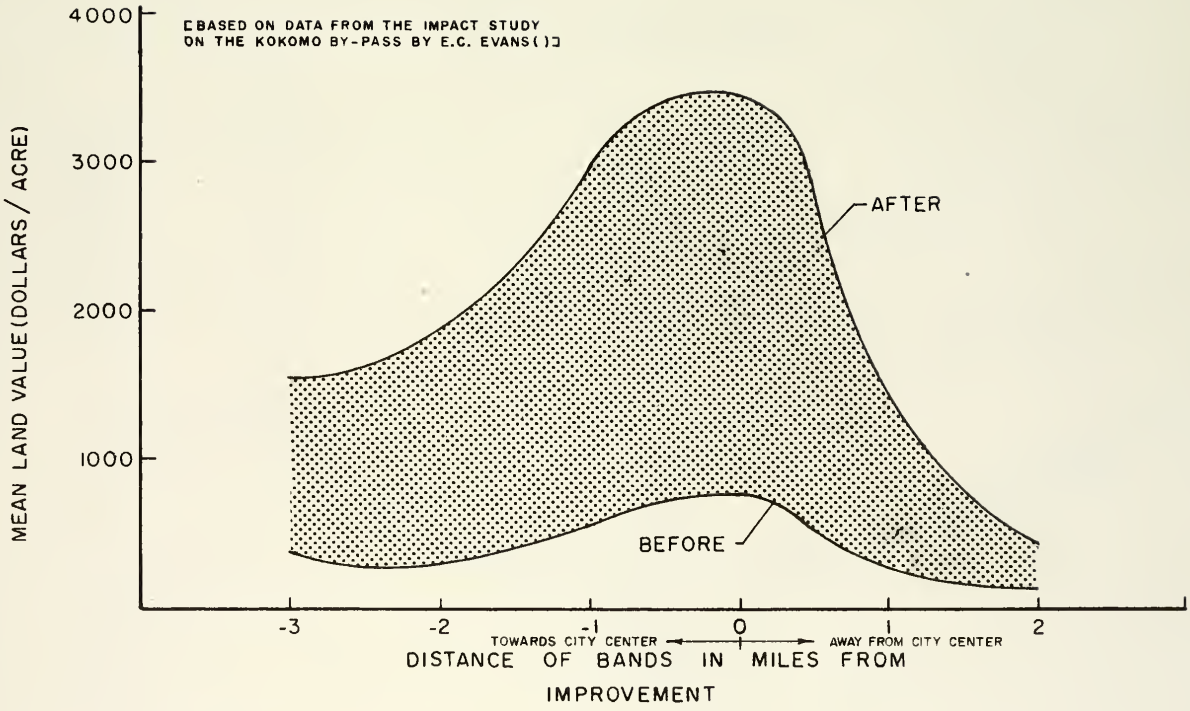



FIGURE 15: MEAN LAND VALUE IN CONSTANT DOLLARS AT VARIOUS DISTANCES FROM THE IMPROVEMENT

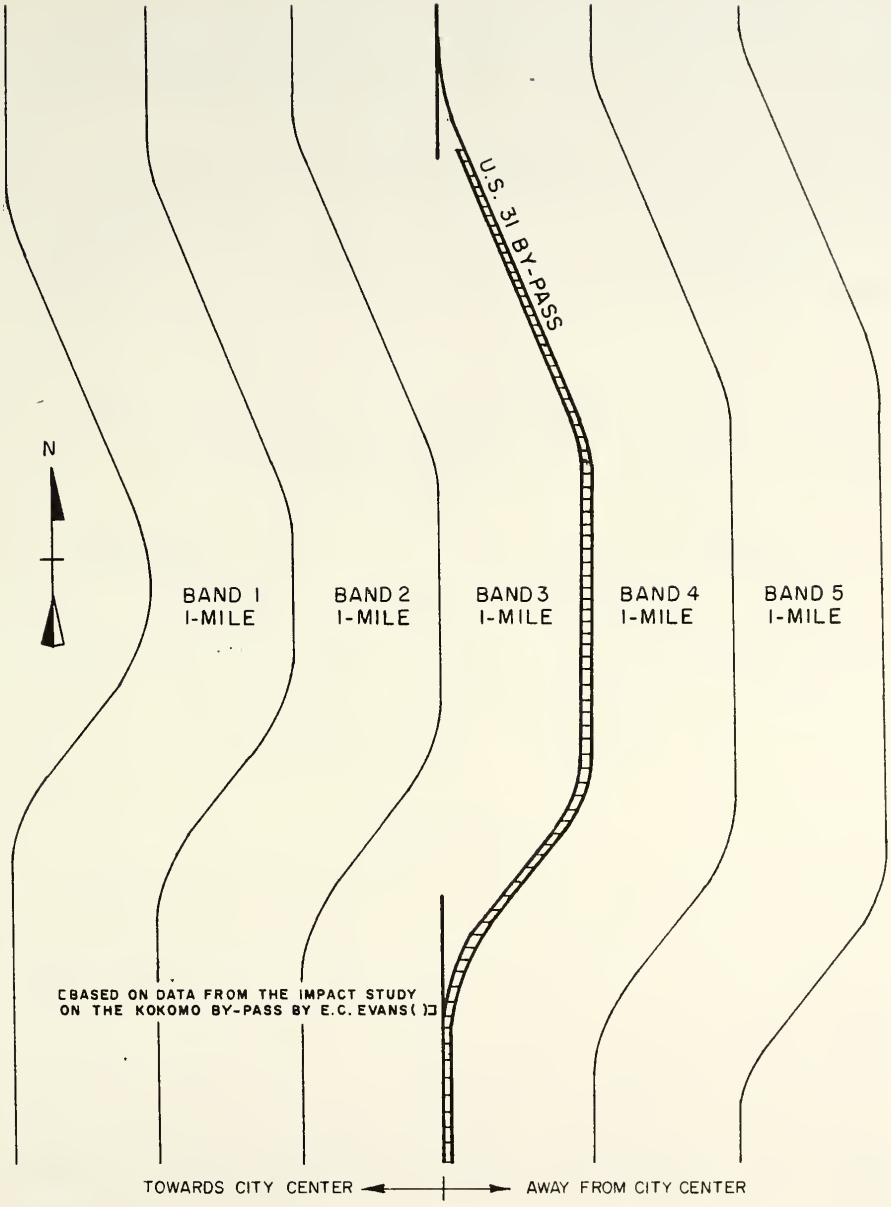



FIGURE 16: LAND VALUE BANDS PARALLEL TO THE BY-PASS

This factor could explain some differences in land values at various distances from the urbanized area.

### Type of Area

This variable is closely related to the preceding variable. It calls for the identification of the area in which the land parcel is located at the time of the construction of the highway improvement. Three types of areas are easily defined:

Urban - denotes that the parcel in question is located in an area that had developed prior to the construction of the highway improvement.

Urban Fringe - denotes that the parcel was in a developing area adjacent to the edge of an urban area.

Rural - denotes that the parcel was in an area that was predominantly used for agricultural purposes and in which no urban fringe development is occurring.

### Time Elapsed

The full effect of a highway improvement on land values might not be indicated during the early period following the construction of the highway improvement. This time lag is required for people to know about the change and to react to it. However a well-publicized highway improvement might start to show its effects on land values even prior to its construction. Figure 17 shows how the value of land parcels along the Ventura Boulevard reacted to various stages of its development.

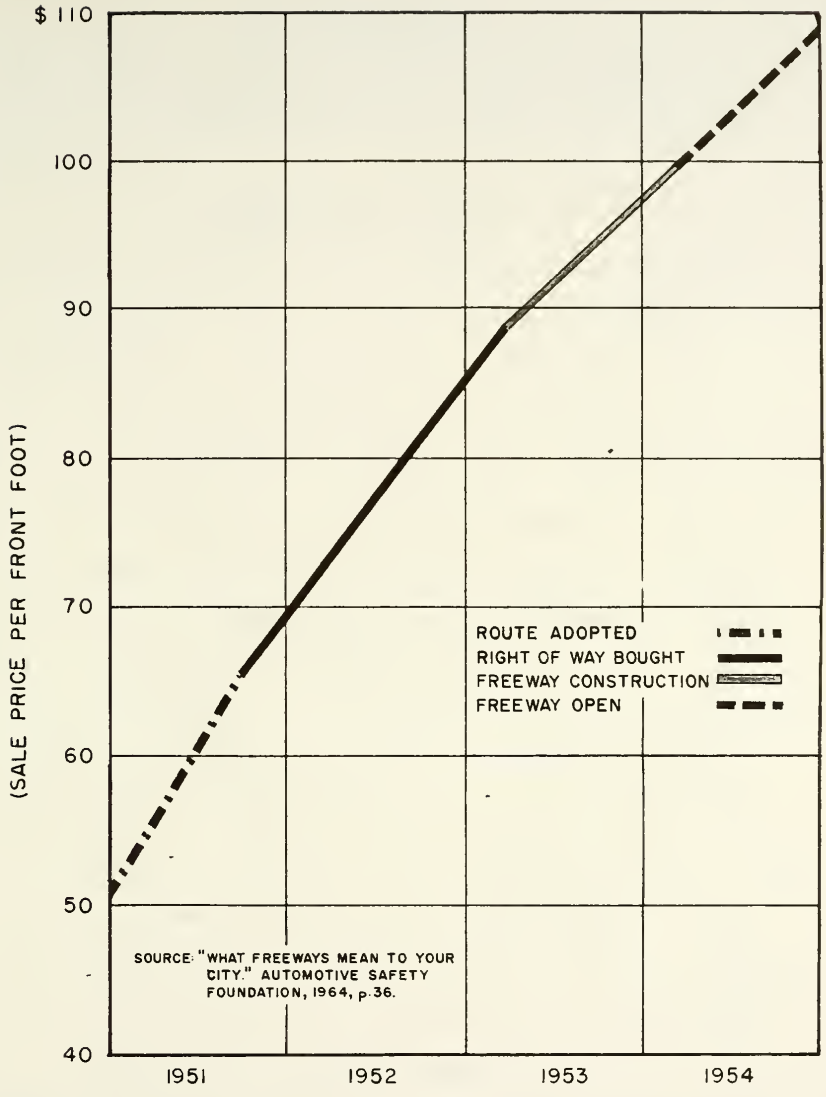



FIGURE 17: LAND PRICE CHANGES ALONG VENTURA BOULEVARD IN CAMARILLO, CALIFORNIA.

## Size of Land Parcel

This variable is included in impact analysis to isolate any differences in the response variable that might be attributable to parcel size.

## Type of Access Control

The type of access control between the parcel and the specific highway improvement for which a portion of the parcel was taken might have some significance of how its proximity to the improvement can be put to full utilization. The inclusion of this variable is therefore useful in inferring the effect of access control on the dependent variable. The types of access control usually defined are as follows:

Full - This designates those parcels separated from the improvement by a fenced right of way and with access to the new highway only via an interchange.

Partial - This type of control indicates the highway improvement being made had only partial control of access (usually some intersections at grade) or the specific improvement for which a portion of the parcel was taken was not fully controlled but an improvement required because of the construction of a nearby fully controlled access highway (for example for the construction of a non-interchange crossing of an interstate highway).

None- This represents the case in which there is little control over access between the parcel and the highway improvement, each property owner has access to the facility.

## Visual Exposure

The number of vehicles per day passing by a land parcel which is adjacent to a highway represents a useful measure of the exposure or the business potential of the parcel in the land market. All available evidence seems to indicate that land adjacent to major highways is a choice location for industries and commercial establishements because of the advertising value received by these firms due to the visual exposure to traffic. Logic seems to suggest that the anticipated future traffic flow might be more indicative of this exposure than existing volumes.

Other variables which measure the exposure of the land parcel from the users of the new facility include visibility of the parcel and its elevation relative to the highway improvement.

# CHAPTER FOUR

# DEVELOPMENT OF A DETERMINISTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM

## Modeling The Real World

In most problem-solving situations the needs often arise to recreate the phenomenon under study, to control the causative factors in a sensitivity analysis and/or to identify the relationship between the response variable and the explanatory variables. Unfortunately however, in many real-life situations, it is not always feasible to satisfy such desires.

A chemist can successfully duplicate a chemical experiment in the laboratory and control the amount of reactants taking part in a chemical reaction. An urban planner or a transportation engineer does not have such a control over the urban area. The task of planning for the urban area would be greatly simplified if it were possible for a transportation planner to observe directly in his laboratory what would happen following any decision he makes affecting the urban network. Even if it were possible to put the urban area in a test-tube the planner would still be unable to introduce into his experiment the exact role played by human elements.

In spite of this lack of perfect control over the forces shaping the urban scene, engineers are called upon by society to make predictions about some real world occurrences and to devise solutions for them. To meet this responsibility problem-solvers often have to resort

to making some approximations as to the real nature of the problem, and on the basis of simplifying assumptions attempt to tackle the problem. These abstractions which enable us to simulate a real world problem underly the modeling process.

Saaty ( 1 ) offered a good definition of a model when he described it as "an objective representation of some aspects of a problem by means of a structure enabling theoretical subjective manipulations aimed at answering certain questions about the problem". The role played by models in analysis is depicted in Figure 18 .

Models can help us to understand the nature of the factors that explain the observed impact of highway improvements on land value. They provide a useful insight into the problem and can be useful in communicating the nature and behavior of engineering creations to those persons who must approve, build, operate and maintain them. Models provide decision-makers with a high-powered predictive tool. They provide engineers with predictions of solution behavior without the necessity of physically testing the solution. The manipulation of mathematical and simulation models through the use of high-speed computers enables the engineer to evaluate several alternative solutions with less time and money. Also in certain situations, such as in the prediction of highway impacts where subjective judgements are not completely reliable and where real life experimentation is prohibitively expensive, prediction models provide the only answer. Models, if properly used, can provide us with a logical and systematic approach towards solving the impact problem.

On the basis of the functions served, Lowry ( 2 ) suggested that any model belongs to one of three classes; descriptive models,

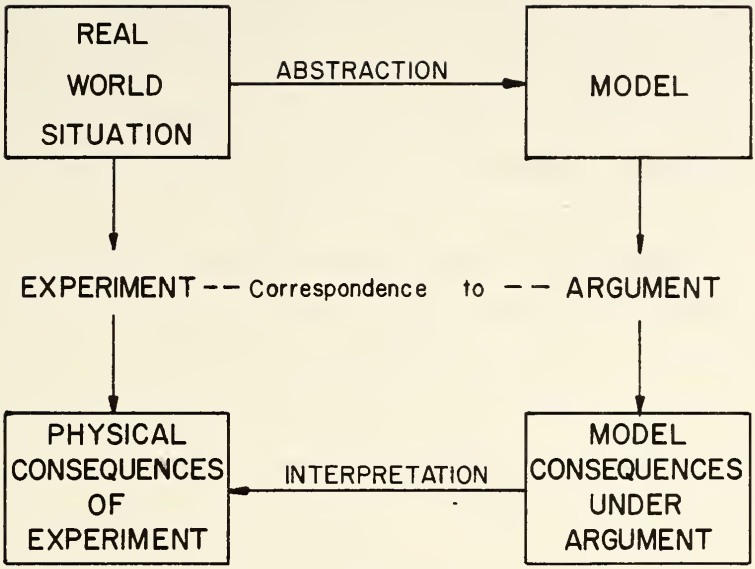



FIGURE 18: THE ROLE OF MODELS IN ANALYSIS

SOURCE: FOUNDATIONS AND TOOLS OF OPERATIONS RESEARCH AND THE MANAGEMENT SCIENCES. UNIVERSITY OF MICHIGAN, SUMMER CONFERENCE NOTES 1962

predictive models and planning models.

Descriptive models attempt to explain the relevant features of the subject-matter. The degree of success attained in reaching this goal is measured by the ratio of the necessary inputs to the possible outputs of the model, the cost and accuracy of the model as compared to the direct observation of the variables under question and the utility achieved in applying the model to times and places other than those from which it was calibrated. While descriptive models provide an understanding of the nature of the problem, they fail to provide planners with the forecasts needed to tackle future problems.

Predictive models go a step beyond the qualitative relationships established between variables by descriptive models. Predictive models provide the planner with a knowledge of the rate of change of the variables with respect to one another. Some of these predictive models handle uncertainties surrounding most decision problems by conditional statements. The analysis of highway impacts is a good example of situations calling for such conditional predictions.

The third group of models is the planning model. Such a model incorporates the notion of conditional prediction and in addition evaluates the outcomes of various alternatives in terms of the prestated goals of the planner.

The above classification of models is based on the uses to which they are put. Models could also be classified on the basis of their relationship to the real world situations which they are supposed to represent. On this basis, models can in general be classified either as iconic, analogue or symbolic models. Iconic models are those that

bear a physical resemblance to their real-life counterparts. They can be such two-dimensional representations as a blueprint or such three-dimensional representations as a toy-train or a replica of the earth. These iconic models serve as an aid to visualization and can be very useful in transmitting information about the subject matter to an observer.

Analogue models are those representations that act like the problem they are designed to represent. A scale model of a dam spillway is an example of an analogue model.

Symbolic models, which are also referred to as mathematical models, utilize symbols and mathematical formulations to describe the real world situations they are designed to represent. The attempt at modeling in this research is directed at developing such mathematical models.

## Regression Theory

Mathematical expressions of the relationships between variables were ushered into the planning process through the desire of planners to obtain a more concise and accurate expression of the forces operating on the urban scene. Through the formulation of mathematical models, planners are able to develop functional relationships between response variables and explanatory variables. In many cases procedures based on regression analysis are utilized. Regression analysis can be simply described as a statistical procedure for developing a mathematical equation which relates two or more variables.

The origin of regression theory is credited to Sir Francis Galton, an English scientist of the 19th century. His interest on

the theory of evolution, especially in relationship to human characteristics, led him to this question:

"If two tall parents tend to have still taller children and two short parents tend to have still shorter children, why has not the human race tended over the centuries to breed giants on the one hand and dwarfs on the other?" (3)

In search of an answer to this question he collected data on the average heights of parents and their fully grown children. On plotting the data he found that the line showing the relationship between average height of parents and the average height of the children was not sloped at 45° but at a smaller angle. That is to say that the heights of children of tall parents were on the average less than the average height of their parents while the heights of children of short parents were on the average taller than the average height of their parents. Galton thus concluded that the height of the offspring tended to "regress" back to the average height of the race. Ever since this finding, lines of average relationships in statistics are called lines of regression.

## Calibration Procedure

The variations in the observed values of the dependent and the independent variables in regression analysis are such that it would be possible to evolve several equations to describe their relationship. Our objective, however, is to obtain the optimum mathematical equation; optimum according to some specified criteria. Two mathematical criteria in use are the "least square" criterion and the "maximum likelihood" criterion. The mathematical and computational procedures

associated with the least square criterion provide the researcher with an equation whose coefficients are such that the sum of the squares of the differences between the observed values of the dependent variable and the values of the dependent variable predicted by the equation will be a minimum. Thus for the assumed functional form relating the dependent and the independent variables the least square line provides the best fit to the data.

The maximum likelihood criterion appears to offer some theoretical advantages over the least square criterion but as the computation required is about four times as long, the least square approach is generally preferred. ( 4 )

The regression approach is designed for the purpose of predicting the response variable if the explanatory variables are given. It does not serve well in general for the estimation of the values of the constant term and the regression coefficients as they exist in the hypothetical infinite population corresponding to the sample ( 5 ). However, if the assumptions stated below are accounted for the regression technique can yield the best prediction of the response variable and also the best estimates of the constant and regression coefficients. Therefore for a defensible application of the regression technique and for the validity of statements made about the statistical significance of measures, the following assumptions must at least be satisfied:

1. Linearity: A linear relationship must exist between the dependent and the independent variables. When linearity is violated the appropriate transformation should be made to correct for this.

2. Normality: The error term $\epsilon_i$ involved in estimating the dependent variable is normally distributed with mean zero and variance $\sigma^2$

3. Homoscedasticity: The homoscedasticity assumption implies that the variance of the error term $\sigma^2$ (also called the residual variance) is a constant, irrespective of the values assumed by the independent variables

4. Independence of the Error term: The error term $\epsilon_i$ is assumed not to be correlated with the independent variables

5. Non-randomness of the Independent variables: For any set of observations only the dependent variable is a random variable and the independent variables are assumed to be fixed or non-random and measured without any error. The set of the dependent variables, corresponding to a given set of values for the independent variables, is normally and independently distributed.

## Statistical Measures

### The Coefficient of Multiple Determination ($R^2$)

This statistic provides a measure of the proportion of the total variance in the dependent variable that is accounted for by the independent variables in the regression equation over that which could be attributed to the mean of the dependent variable alone.

Considering the two dimensional case, Figure 19 shows that some of the total variation of observations about their mean can be accounted for by the regression line while some still remains unexplained because not all the observations lie on the regression line. Therefore we can say that:

$$\left\{\begin{array}{c}\text{Sum of squares} \\ \text{about mean}\end{array}\right\} = \left\{\begin{array}{c}\text{Sum of squares} \\ \text{about regression}\end{array}\right\} + \left\{\begin{array}{c}\text{Sum of squares} \\ \text{due to regression}\end{array}\right\}$$

$$R^2 = \frac{(\text{Sum of squares due to regression})}{(\text{Total sum of squares about mean})}$$

When all the observations lie on the regression line, the sum of squares about the regression line would be zero because the regression line, in such a case, perfectly explains the observations. This means, in other words, that the sum of squares about the mean is equal to the sum of squares due to regression and $R^2$ takes a value of one. At the other extreme, there is a complete lack of fit and none of the variance is explained by the independent variables included in the model. Then the sum of squares about the mean is equal to the sum of squares about the regression line and $R^2$ is zero.

This statistical measure ($R^2$) thus gives the analyst an idea of

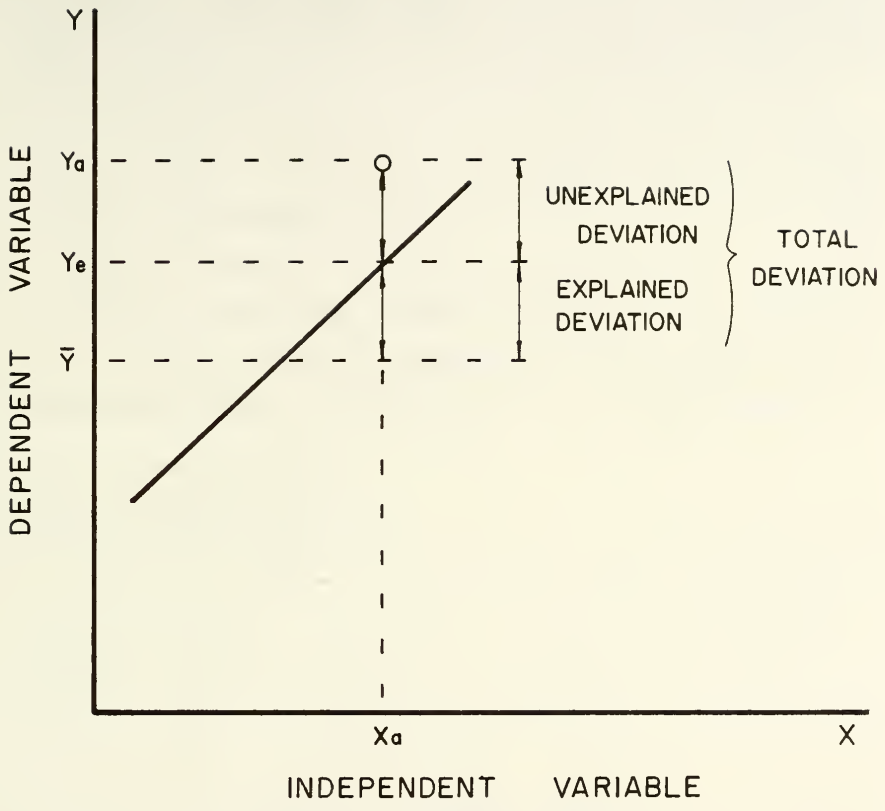



FIGURE 19: TWO DIMENSIONAL REGRESSION

how well the regression equation is doing in explaining the relationship between the independent variables in the equation and the dependent variable. Although a high $R^2$ is desired, it should be interpreted with caution. A high $R^2$ value might indicate we have a statistically reliable regression model but it does not necessarily mean that there is a causal relationship between the explanatory variables and the response variable. It is therefore essential that only the independent variables that are logically related to the dependent variable at hand be allowed to enter into the model.

The square root of the coefficient of multiple determination is R, the coefficient of multiple correlation. This is a measure of the association between the independent variables in the equation and the dependent variable.

Partial F or Sequential F test

The partial F-test permits the analyst to check at each step of a stepwise regression if the increase in $R^2$ contributed by each added independent variable in the equation is significantly different from zero. The value of the computed partial F statistic is given by:

$$F_{(k,\ n-k-1)} = \frac{(R_k^2/k)}{(1-R_k^2)\ /\ (n-k-1)}$$

where

$n$ = number of observation

$k$ = number of independent variables

$R_k^2$ = coefficient of multiple determination of model with k independent variables.

The F statistic calculated by the above equation is compared to a tabulated $F_{(k,n-k-1,\ 1-\alpha)}$ where $\alpha$ is the probability of type I error or the level of significance.

The t Statistic

To test the hypothesis that each of the estimated regression coefficients is significantly different from zero the t-statistic is often used. A regression coefficient which fails to satisfy the t test can be regarded insignificant in contributing to the observed response; and should be deleted from the equation. This statistic is expressed by the ratio of each regression coefficient to its standard error. The square of the t-statistic has the same distribution as the F-statistic. Thus the two-tailed t test is equivalent to the one tailed F-test.

The Standard Error of Estimate

The standard error of estimate is a measure of the spread of the observed data points about the regression line. It therefore provides a measure of the expected error involved in predicting the dependent variable with the established equation. Therefore the smaller the value of the standard error of estimate, the more precise the predictions would be. However, it should be recognized that this statistic refers only to the data used in fitting the regression equation. It does not provide an estimate of the probable error in predicting a value of the response variable for those values of the explanatory variables that are not covered by the data points used in calibrating the model. This type of estimate of probable error for future values of the dependent variable is given by the "standard

error of forecast".

The standard error of estimate is similar to the standard deviation of a set of data about its mean value. We can also build confidence bands about the regression line in a way similar to those that can be set about the mean. Then we can also assert that 68.3% of the observations can be expected to fall within plus or minus one standard error from the regression line, 95% fall with plus or minus two standard errors or that 99.7% of all observations are within plus or minus three standard errors of the regression line.

## Uses of Dummy Variables In Regression Analysis

In modeling, some of the independent variables relevant to the problem at hand cannot be measured on a continuous scale. Cross classification provides one way of analyzing such variables. This procedure does not require an assumption as to the distribution of the variables. Unfortunately however, a large sample size is required to reach any worthwhile conclusion. Ezekiel and Fox summarized these shortcomings of the cross classification technique in these words:

"----it provides no measure of how important the relation shown is as a cause of variation in the factor being studied, or of how closely that factor may be estimated from the others on the basis of the relations shown---. Cross-classification does not determine the relationships where many variables are involved so satisfactorily as does multiple regression".( 6, p.394)

An alternative analytical procedure that also deals with this class

of variables is the dummy variable technique. The use of dummy variables provides the analyst with a simple method of handling discontinuous variables. Categorical variables such as race, sex, occupation and region can then be included in regression analysis. An article by Daniel B. Suits ( 7 ), then of the University of Michigan was perhaps the first in the literature of statistics on the use of dummy variables in regression analysis.

The dummy variable technique possesses some distinct advantages which ought to be recognized. As mentioned earlier, it provides a method of including categorical data into regression analysis. Also unlike the cross classification procedure, it does not require a large number of observations. The dummy variable technique can also be applied to introduce quantitative independent variables into a regression model. These quantitative variables are transformed into a dummy format by stratifying them into a number of discrete classes, each indicating a range within which values of the original variable falls. A dummy variable is then assigned to each of the discrete classes. By indicating only the dummy class into which an observation falls, errors made in reporting and recording data on the explanatory variables become less disturbing as to their effect on the response variable. For instance, if a dummy variable is used to represent land parcels whose sizes are within the range of 50 to 150 acres it becomes insignificant if a parcel whose size is 100 acres is reported as 120 acres. The dummy variable technique also enables the researcher to eliminate the biases that are usually involved in assuming the form of functional relationship between the variables. This is so because it allows the identification

of nonlinear relationships without any prior assumption as to the character of nonlinearity. Thus assumptions such as the linearity assumption do not have to be made. To test for linear relationships between the response and quantitative explanatory variables, the dummy regression coefficients ($b'_{jk}$) are plotted against corresponding class values of the parent independent variable. If a linear plot is obtained then the linear assumption is substantiated. Otherwise a non-linear relationship exists.

## The Dummy Variable Technique

In this technique a value of one is assigned to the dummy variable corresponding to the class within which the value of the observed independent variable falls; all other dummy variables take on the value of zero.

The estimating equation generally takes the form

$$\hat{Y} = b_o + \sum_{i=1}^{M_1} b_i X_i + \sum_{k=1}^{M_2} \sum_{j=1}^{M_3} b_{jk} Z_{jk} \quad \text{------(4-1)}$$

where:

$Y$ = the estimate of the dependent variable

$b_o$ = regression constant

$b_i$ = regression coefficient for the $i^{th}$ linear dependent variable denoted by $X_i$

$b_{jk}$ = coefficient of the $j^{th}$ dummy variable in the $k^{th}$ dummy variable set

$Z_{jk}$ = 1 if observed variable falls into the $j^{th}$ class and 0 if not in $j^{th}$ class

$M_1$ = total number of linear variables

$M_2$ = total number of dummy variable sets

$M_3$ = number of classes within a dummy variable set.

The least square criterion can then be used to evolve a regression surface where the sum of squared deviations of all the observations about it is a minimum. Solution to the developed "normal" equations yields the desired estimates of the constant term and the regression coefficients. However the introduction of qualitative variables via the use of dummy variables into the regression model renders the normal equations indeterminate. This indeterminancy is present because there are more coefficients than there are independent normal equations based on the least squares criterion. Two constraints are generally imposed to remove the existing indeterminancy.

First Type of Constraint

The first approach constrains one of the coefficients in each of the groups to equal zero. For an illustration of this constraint let us assume a simplified case of the impact problem in which the response variable is Y and the independent variables are size of land parcel (a linear variable), areal location and type of land use. Suppose the last two independent variables have four classes each. Then the estimating equation is:

$$\hat{Y} = b_o + b_1 X_1 + \sum_{k=1}^{2} \sum_{j=1}^{4} b_{jk} Z_{jk} \qquad (4\text{-}2)$$

As stated earlier, it would be technically impossible to estimate the above equation directly. The inherent indeterminancy can be eliminated if one of the coefficients in each of the two dummy variable sets is preassigned a value of zero. Therefore, arbitrarily set

$$b_{41} = 0 \text{ and } b_{42} = 0$$

This constraint thus reduces the number of coefficients to be determined by $M_2$ (in this example $M_2 = 2$). The reduced estimating equation is now:

$$\hat{Y} = b'_o + b'_1 X_1 + \sum_{k=1}^{2} \sum_{j=1}^{3} b'_{jk} Z_{jk} \qquad (4\text{-}3)$$

Now each estimated $b_{jk}$ measures the net effect that membership in the $j^{th}$ dummy class has on the dependent variable as compared to membership in the excluded classes.

Second Type of Constraint

In using the first type of constraint, all estimated regression coefficients are interpreted with reference to that of the omitted dummy class. It might be more desirable to obtain regression coefficients which indicate differences from the overall mean of the dependent variable rather than from the base classes.

The type of constraint to be introduced next leads to the estimation of coefficients of this kind. To achieve this, the weighted sum of the coefficients in each dummy set representing a single factor is set equal to zero ( 8 ).

Applying this constraint to our example we have:

$$\sum_{j=1}^{4} b^{*}_{j1} z_{j1} = 0 \qquad (4\text{-}4)$$

$$\sum_{j=1}^{4} b^{*}_{j2} z_{j2} = 0 \qquad (4\text{-}5)$$

where

$z_{jk}$ = number of parcels in dummy class j of the $k^{th}$ dummy set

Solving for $b^{*}_{41}$ and $b^{*}_{42}$ from equations 4 and 5 respectively we obtain

$$b^{*}_{41} = -\left\{\frac{1}{z_{41}}\left(\sum_{j=1}^{3} b_{j1} z_{j1}\right)\right\} \qquad (4\text{-}6)$$

$$b^{*}_{42} = -\left\{\frac{1}{z_{42}}\left(\sum_{j=1}^{3} b_{j2} z_{j2}\right)\right\} \qquad (4\text{-}7)$$

Substituting for $b^{*}_{41}$ and $b^{*}_{42}$ from equations (4-6) and (4-7) in equation (4-2) we get:

$$\hat{Y} = b^{*}_{o} + b^{*}_{1}X_{1} + \sum_{k=1}^{2}\sum_{j=1}^{3} b^{*}_{jk}\left(Z_{jk} - \frac{z_{jk}}{z_{4k}} Z_{4k}\right) \qquad (4\text{-}8)$$

Now if we define new variables

$$Z^{*}_{jk} = Z_{jk} - \left(\frac{z_{jk}}{z_{4k}}\right) Z_{4k} \qquad (4\text{-}9)$$

we can rewrite equation (4-8) as

$$Y = b_o^* + b_1^* X_1 + \sum_{k=1}^{2} \sum_{j=1}^{3} b_{jk}^* Z_{jk}^* \qquad (4\text{-}10)$$

The above equation can be solved using the usual least square approach to obtain estimates of the regression coefficients. The resulting values of the regression coefficients are then substituted into equations (4-6 and 4-7) to determine the remaining regression coefficients.

Examining equation (4-10) it is evident that each of the new independent variables $Z_{jk}^*$ has a mean of zero. This infers that the constant term $b_o^*$ possesses a value equal to that of the mean of the dependent variable. Therefore each estimate of regression coefficient for a dummy class is the net difference from the grand mean which is attributable to membership in that dummy class.

Although the two types of constraints discussed above produce different regression coefficients when they are each imposed on the normal equations, they are both effective in surfacing the differences among the coefficients. The only dissimilarity is the basis of reference used by each type of constraint to measure this difference. Therefore we can transform the results from the use of the first constraint to those obtained from the second type of constraint. The similarity in the final equations (4-3) and (4-10) developed by using the first and second constraints respectively.is of some importance. The only difference is in the magnitudes of the constant term and the various regression coefficients. We can therefore see that each of

the coefficients $b_{jk}$, $b^{*}_{jk}$ developed from using the first and second constraints respectively differ by a constant, C say. Therefore we write:

$$b^{*}_{jk} = b'_{jk} + C$$

Returning to our sample situation we have

$$b^{*}_{jl} = b'_{jl} + C$$

Applying the first constraint in which we set the coefficient of one of the dummy classes equal to zero. (In this case $b'_{4l}$) we have

$$b^{*}_{4l} = b'_{4l} + C$$
$$= 0 + C$$
$$b^{*}_{4l} = C$$

Now applying the second constraint

$$\sum_{j=1}^{4} (b'_{jl} + C)\, z_{jl} = 0$$

$$C = -\left\{ \frac{\sum_{j}^{3} b'_{jl}\, z_{jl}}{\sum_{j=1}^{4} z_{jl}} \right\}$$

$$b^{*}_{jl} = b'_{jl} - \left\{ \frac{\sum_{j=1}^{3} b'_{jl}\, z_{jl}}{\sum_{j=1}^{4} z_{jl}} \right\}$$

$$\text{i.e} \quad b^{*}_{j1} = b'_{j1} - \sum_{j=1}^{4} P_{j1}\, b'_{j1}$$

where

$$P_{j1} = \frac{z_{j1}}{\sum_{j=1}^{4} z_{j1}}$$

Therefore in general we can write

$$b^{*}_{jk} = b'_{jk} - \sum_{j=1}^{M_3} P_{jk}\, b'_{jk}$$

where

$b^{*}_{jk}$ = the adjusted regression coefficient of the $j^{th}$ dummy variable in the $k^{th}$ dummy variable set, (indicating the influence of the $j^{th}$ dummy class in the $k^{th}$ dummy variable set on the response variable <u>with respect to the overall mean</u>)

$b'_{jk}$ = the uncorrected regression coefficient (indicating the relative influence of the $j^{th}$ dummy class in the $k^{th}$ dummy variable set on the response

variable as compared to that of the omitted dummy class in the set)

$P_{jk}$ = is a weighting factor indicating the fraction of the sample in the $j^{th}$ dummy class in the $k^{th}$ dummy variable set

A Geometrical Interpretation

The various implications accruing from the use of dummy variables in regression analysis can be viewed more intimately through an example. Suppose we are interested in measuring the response on land value following a given highway improvement; an analysis of the situation reveals that the estimating equation is

$$Y = 22.5 - 5.0X - 2.5Z_1 - 1.5Z_2$$

where

Y = estimated change in unit land value expressed in constant dollars

X = distance from the land parcel to the highway improvement in feet

$Z_1$ , $Z_2$, = Dummy variables representing two of the three classes of a qualitative variable, type of area.

A geometrical representation of the estimating equation is as shown in Figure 20 . The regression coefficients in the equation are estimated with the omitted class ($Z_3$) as a base. The influences

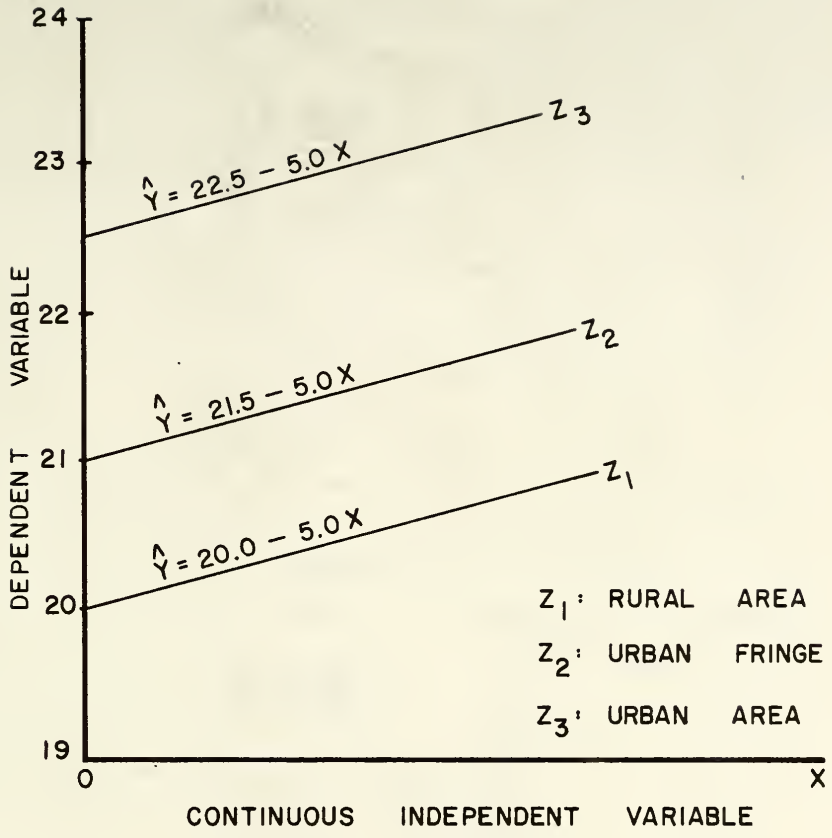



FIGURE 20: EFFECT OF MEMBERSHIP IN A DUMMY VARIABLE CLASS ON REGRESSION LINE.

introduced by the different classes in the dummy variable set are exhibited by the shifts in the intercepts of the regression of Y on X with area $Z_3$ taken as a base. From these intercepts we can infer for example that parcels in rural areas, all other things remaining unchanged, experience a change in land value which is two and half dollars per acre less than similar parcels located in an urban area. Similarly the differences in the intercepts of the regression lines for classes 2 and 3 reveal that an acre of urban land is enhanced by the highway improvement by one and half dollars more than an acre in the urban fringe.

Measures of Importance

As noted earlier, the coefficient of multiple determination $R^2$ provides the analyst with a measure of how well the evolved regression equation is doing in explaining the variations in the response available. An insight into the proportion of the total variation in the response variable that is explained by any given factor, such as A, can be determined by the partial $R^2$ defined as:

$$\text{Partial } R^2 = \frac{R_A^2 - R_B^2}{1 - R_B^2}$$

Where

$R_A^2$ = coefficient of multiple determination for the equation that includes factor A

$R_B^2$ = coefficient of multiple determination before the variables representing factor A were added.

In addition to knowing that what proportion of the total variation is explained by factor A it might also be useful to know whether this net contribution is of any statistical significance. The F ratio as defined below provides the required statistical measure

$$F = \frac{(R_A^2 - R_B^2)\ (N - K_A - K_B - 1)}{(1 - R_A^2)\ (K_A)}$$

Where

$K_A$ = number of independent variables representing factor A

$K_B$ = number of independent variables other than those representing factor A

$N$ = number of observations

$R_A^2$ = coefficient of multiple determination for equation with ($K_A = K_B$) variables

$R_B^2$ = coefficient of multiple determination for equation with $K_B$ variables.

Relative Importance of Each Dummy Variable Set.

In situations in which linear quantitative independent variables $X_i$ are involved, a direct comparison of the relative importance of each variable in the regression equation is only possible when regression coefficients are in the same units. To achieve a non-dimensional basis for comparison, each coefficient is multiplied by the ratio of the standard deviation of the linear variable to the standard deviation of the dependent variable. The final product is the beta coefficient

$\beta$ for that variable.

For the $K^{th}$ set of dummy variables the beta coefficient $\beta_k$ is given by:

$$\beta_K = \frac{1}{S_Y}\left(\frac{\sum_{j=1}^{C} z_{jk}\, b^{*2}_{jk}}{\sum_{j=1}^{C} z_{jk}}\right)^{1/2}$$

$$= \frac{1}{S_Y}\left(\sum_{j=1}^{C} P_{jk}\, b^{*2}_{jk}\right)^{1/2}$$

Where

$S_Y$ = standard deviation of the dependent variable, Y.

Testing for Significance

The hypothesis to be tested here is whether or not the regression coefficient developed for each dummy class in a given dummy variable set is different from zero. In testing whether the difference between a given dummy class and an omitted class in a dummy variable set is significant, the t-test can be used in the usual manner on the corresponding regression coefficient computed by using the first type of constraint. However, if a t-test is to be conducted for significance for all the dummy variables representing all the classes in a given dummy variable set, the test should be carried out on the adjusted regression coefficients ($b^{*}_{jk}$). This is another reason for adjusting regression coefficients.

# CHAPTER FIVE

# DESIGN OF STUDY

## Methodology

The method employed in the model evolution phase of this research is that of multivariate linear regression analysis. Two sets of models were developed; these are identified as the I-set and the IF-set. The I-set used data from the state of Indiana, and the IF-set used data from the states of Indiana and Florida. Each of these sets contains ten models, identified in Figure 21 .

The general model is of the form:

$$Y = \beta_o + \sum \beta_i \alpha_i + \sum \beta_j \gamma_j + \epsilon$$

where

$Y$ =dependent variable

$\alpha_i$ =linear independent variable

$\gamma_j$ =dummy independent variable

$\beta_o$ =constant term

$\beta_i, \beta_j$ =regression coefficients

$\epsilon$ = error term or the net effect of contributions to the value of the dependent variable attributable to other variables not included in the model.

To calibrate this assumed functional form, the empirical regression coefficients $b_o$, $b_1$,....$b_n$ (estimates of $\beta_o$, $\beta_1$,....$\beta_n$ respectively)

| MODEL NAME | DEPENDENT VARIABLES | | | | IVE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | DIFFERENCE IN CONSTANT DOLLARS (DDIFF) | PERCENT DIFFERENCE (PDIFF) | RECOVERY RATE (RVALUE) | SIZE $X_4$ | TYPE OF AREA | | | TYPE OF ACCESS | | |
| | | | | | URBAN $Z_{19}$ | FRINGE $Z_{20}$ | RURAL $Z_{21}$ | FULL $Z_{22}$ | PART. $Z_{23}$ | NONE $Z_{24}$ |
| CC-1 | ● | | | ● | ● | ● | ● | ● | ● | ● |
| CC-2 | ● | | | | ● | ● | ● | ● | ● | ● |
| CC-3 | ● | | | | ● | ● | ● | ● | ● | ● |
| PC-1 | | ● | | ● | ● | ● | ● | ● | ● | ● |
| PC-2 | | ● | | | ● | ● | ● | ● | ● | ● |
| PC-3 | | ● | | | ● | ● | ● | ● | ● | ● |
| RR-1 | | | ● | ● | ● | ● | ● | ● | ● | ● |
| RR-2 | | | ● | | ● | ● | ● | ● | ● | ● |
| RR-3 | | | ● | | ● | ● | ● | ● | ● | ● |
| RR2A | | | * | | ● | ● | ● | ● | ● | ● |

* LOG(RV

FIGURE 21:

| MODEL NAME | DEPENDENT VARIABLES | | | INDEPENDENT VARIABLES | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | QUANTITATIVE | | | | | | | | | | | | | QUALITATIVE | | | | | | | | | | | | | |
| | DIFFERENCE IN CONSTANT DOLLARS (DDIFF) | PERCENT DIFFERENCE (PDIFF) | RECOVERY RATE (RVALUE) | SIZE $X_4$ | TIME $X_5$ | LOG (SIZE) | LOG (TIME) | SIZE | | | TIME | | | HIGHWAY TYPE | | | LAND USE CATEGORY | | | | TYPE OF AREA | | | TYPE OF ACCESS | | | | | |
| | | | | | | | | SIZEA $Z_6$ | SIZEB $Z_7$ | SIZEC $Z_8$ | TIMEA $Z_9$ | TIMEB $Z_{10}$ | TIMEC $Z_{11}$ | INTERS. $Z_{12}$ | PRIM. $Z_{13}$ | SEC. $Z_{14}$ | RES. $Z_{15}$ | COMM. $Z_{16}$ | AGRIC. $Z_{17}$ | VAC. $Z_{18}$ | URBAN $Z_{19}$ | FRINGE $Z_{20}$ | RURAL $Z_{21}$ | FULL $Z_{22}$ | PART. $Z_{23}$ | NONE $Z_{24}$ | | | |
| CC-1 | ● | | | ● | ● | | | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| CC-2 | ● | | | | | ● | ● | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| CC-3 | ● | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| PC-1 | | ● | | ● | ● | | | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| PC-2 | | ● | | | | ● | ● | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| PC-3 | | ● | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| RR-1 | | | ● | ● | ● | | | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| RR-2 | | | ● | | | ● | ● | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| RR-3 | | | ● | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |
| RR2A | | | * | | | ● | ● | | | | | | | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | | |

* LOG(RVALUE)

FIGURE 21: IDENTIFICATION SYSTEM FOR DEVELOPED MODELS

were determined from sets of observations using the least square criterion.

The computations leading to the determination of the values of these regression coefficients were performed on the CDC 6500 computer at Purdue University using the stepwise regression program (BMD 02R) developed at the University of California at Los Angeles ( 1 ). In the procedure outlined in this program, an "F-to-remove" value for each variable in the regression equation and "F-to-enter" value for each independent variable not in the equation are computed at each step. The variable added at each step is the one which reduces the residual sum of squared deviations by the largest amount. Independent variables are added or deleted in accordance with the following criteria:

1. If there are one or more independent variables in the regression equation whose control value as indicated on the control-delete card is 2 (i.e. it is a free variable) and whose F value is less than the "F-to-remove" value specified on the sub-problem card the one with the smallest F-value will be removed.

2. Assuming no variable is removed by the first criterion, and that there are one or more independent variables not in the regression equation which pass the tolerance test and have control values of 3 or more (i.e. forced variables), the one which has the highest control value and the highest F-value among all others, with the same control value, will be added.

Any independent variable not in the regression equation is said to pass the tolerance test if its tolerance value exceeds or equals the

"minimum tolerance value" specified on the sub-problem card.

3. If no variable is removed by the first criterion or added by the second and there are one or more independent variables not in the regression equation which pass the tolerance test, have a control value of 2 (i.e. a free variable) and an F-value greater than or equal to the "F-to-enter" value specified on the sub-problem card, the one with the highest F-value will be added.

In this research the values of 0.005, 0.01 and 0.001 were specified for "F-to-remove" "F-to-enter" and tolerance values respectively. The stepwise procedure was terminated when no variable was added or removed by any of the above three criteria.

## Variables Used in the Models

The complete set of variables both dependent and independent used in constructing the several multiple regression models in this research and shown in Figure 21 are further described in this section.

### Dependent Variables

Fluctuations in land values were selected as indicators of what happens to a land parcel when the road situation changes. Three measures of these fluctuations which were used in this research as dependent or response variables are as follows:

1. Change in land value in constant dollars (DDIFF)
2. The percent change in land value (PDIFF)
3. The recovery rate (RVALUE)

Using the first measure, the response variable DDIFF could be expressed as:

$$Y_H = Y_T - Y_G$$

or

$$Y_H = \left\{LV_A - LV_B\right\} - Y_G$$

Where

$Y_H$ = DDIFF = change in land value attributable to the change in the highway system

$Y_H$ = total observed change in land value

$Y_G$ = general change in the land market in the area attributable to other causes apart from the highway improvement

$LV_B$ = land value in constant dollars before the opening of the highway improvement

$LV_A$ = land value in constant dollars after the opening of the highway improvement.

The magnitude of $Y_G$ is often estimated by observing the change in the land value of a control parcel during the "before" and "after" periods. Such a control parcel is assumed to be unaffected by the highway improvement. Otherwise it is very much similar to the parcels under study.

Researchers at Purdue University have spotlighted the use of the band concept in the estimation of $Y_G$. The use of the control approach is based on the assumption that the selected control area is not affected

by the improvement. The selection of such an area is difficult. Although the control area is said to be sufficiently far away from the improvement that it can be assumed unaffected, it remains a subjective matter to determine how far is far. The band concept overcomes this deficiency by tracing the change in land value at various distances from the improvement. The change in land value with distance appears as a decaying function. After some distance the curve showing this relationship starts to level off and remains practically constant with increasing distance. The magnitude of the change where this leveling off occurs provides an estimation of $Y_G$.

The second measure of the response variable is of the same form as the first except that the change is expressed as a percentage of the before value. The second measure takes the form:

$$P_H = \left\{ \frac{LV_A - LV_B}{LV_B} \right\} 100 - P_G$$

where

$P_H$ = PDIFF - percent difference in land value attributable to the improvement.

$P_G$ = general percentage change in the land market in the area not attributable to the highway improvement.

The third measure of the response variable as used on this study is defined as follows:

$$R = \left( \frac{P_A}{P_B} \right) \times 100$$

where

$R$ = RVALUE = recovery rate expressed as a percentage

$P_A$ = unit price of remainder sold (dollars per acre)

$P_B$ = unit price of original parcel before severance

By this definition $R > 100$ implies that the remainder has increased in value. By the same token $R < 100$ means that the remainder has declined in value. It should be noted that this type of comparison does not consider any payment that might have been made for damages.

## Independent Variables

The following independent variables were used in this research:

1. Size of parcel (acres)
2. Time elapsed between completion of highway improvement and sale of parcel (months)
3. Type of highway improvement (interstate, primary or secondary)
4. Type of land use (residential, commercial, agricultural or vacant)
5. Type of area (urban, urban fringe or rural)
6. Type of access control (full, partial or none).

Each of the types under variables three to six was treated as a dummy variable which assumed a value of one or zero depending on whether or not it was observed for the parcel in question.

## Data

### Source of Data

The data used in this research were obtained from two sources - the States of Florida and Indiana.

The Florida data were made available to this study through the cooperation of the Bureau of Public Roads and are documented in two reports ( 2 , 3 ). These reports were the results of two studies conducted by the Florida State Road Department in cooperation with the Bureau of Public Roads. The first of these studies examined about one hundred and twenty-eight cases of urban partial takings in Miami, Jacksonville, Tampa, Orlando and Pensacola along Interstate Routes I-4, I-10, I-110, I-95, and I-195. The second study investigated about one hundred and six cases of rural land in Alachua, Baker, Columbia, Duval, Escambia, Hamilton, Hillsborough, Osceola, Polk and Volusia counties along Interstate Highways I-10, I-75, I-110 and I-4. To remove biases that might have existed, these studies excluded sales of properties to relatives of the owners and those that involved inter-corporate transfers. From each of these two Florida studies, a random sample of fifty (50) parcels was selected for use in the research. One shortcoming of these data was the absence of parcels in urban fringe areas.

Data from the State of Indiana were obtained from the work of Stover ( 4 ). In his study, Stover collected data on land parcels selected by a random sample of all the right-of-way projects placed under contract by the Indiana State Highway Commission between

January 1, 1955 and December 31, 1961. The selected highway projects were chosen from two random populations - ninety-nine interstate projects and four hundred and thirty projects on primary and secondary routes. In all sixty-six projects were chosen; thirty-one from the first population and the rest (35) from the second population. A large proportion of interstate projects was included in this selection because they were of special interest at that time. From the sixty-six (66) Indiana study parcels, only thirty (30) had sufficient data on all selected variables to permit their use in the modeling aspects of this research.

## Procedure for Obtaining Land Value Data

Changes in land value were determined in Indiana by obtaining information on the sale of each parcel of land within the effected area. The steps taken in gathering this information were in general as follows:

1. A random sample of the land parcels to be studied was first obtained
2. For each land parcel information as to its location with respect to section and county and with reference to a county map was obtained.
3. Names of land-owners were next obtained from the plat books (usually available from the County Auditors office)
4. History of transfer of each land parcel from one owner to another was traced through time from the transfer books. Usually a set of transfer books was available for each township and each

city within each county. Such pertinent data as date of transfer, date of deed, sale value and acreage were recorded from the transfer books. Only value of the land (without improvements) was obtained.

5. Finally the grantee and/or grantor was interviewed. The grantee was generally contacted first because he was usually easier to locate.

In any instance where the sale price was not recorded, an estimate of the sale price was obtained from the state gross income tax declaration or from the amount of United States Revenue Stamps attached. These stamps and the tax declaration were also used to check the validity of each recorded sale price.

Procedures used in Florida for the collection of their data were similar to those discussed for Indiana.

## Data Collected

The data obtained from the States of Florida and Indiana for use in this research are summarized in Tables 4, 3 . In these tables the following abbreviations were used in recording the desired information under each column.

| Column | Abbreviations |
|---|---|
| Highway Type: | |
| Interstate | I |
| Primary | P |
| Secondary | S |

| | |
|---|---|
| Land use type: | |
| Residential | R |
| Commercial | C |
| Agricultural | A |
| Vacant | V |
| Type of area: | |
| Urban | U |
| Urban Fringe | F |
| Rural | R |
| Type of access control: | |
| Full | F |
| Partial | P |
| None | N |

The unit land values recorded under columns two and three of these tables were computed in constant dollars using 1960 as the base year.

# TABLE 3: INDIANA DATA

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| I1 | 434.00 | 349.00 | 49.130 | 32 | I | A | R | F |
| I2 | 1,000.00 | 1,230.00 | 0.610 | 5 | P | A | F | F |
| I3 | 3,600.00 | 32,400.00 | 0.600 | 8 | P | C | F | N |
| I4 | 1,060.00 | 9,063.00 | 0.410 | 1 | S | R | R | N |
| I5 | 1,960.00 | 11,480.00 | 3.046 | 4 | P | R | F | P |
| I6 | 400.00 | 17,150.00 | 2.000 | 13 | P | A | F | P |
| I7 | 2,615.00 | 552.00 | 1.630 | 4 | P | R | R | P |
| I8 | 408.00 | 238.00 | 18.700 | 9 | I | A | R | F |
| I9 | 1,488.00 | 54,880.00 | 0.223 | 66 | S | R | U | N |
| I10 | 1,530.00 | 4,386.00 | 0.977 | 16 | I | R | F | N |
| I11 | 382.00 | 100.00 | 19.600 | 30 | I | A | R | F |
| I12 | 95.00 | 37.00 | 31.500 | 47 | I | A | R | F |
| I13 | 153.00 | 85.00 | 40.250 | 43 | I | A | R | F |
| I14 | 361.00 | 190.00 | 8.020 | 8 | I | A | R | F |
| I15 | 342.00 | 486.00 | 49.000 | 10 | I | A | R | F |
| I16 | 340.00 | 212.00 | 233.170 | 2 | I | A | R | F |
| I17 | 3,260.00 | 875.00 | 1.830 | 3 | I | R | F | P |

# TABLE 3: INDIANA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| I18 | 1,500.00 | 2,323.00 | 0.260 | 7 | I | R | F | F |
| I19 | 2,078.00 | 98,500.00 | 0.132 | 33 | S | R | U | P |
| I20 | 5,450.00 | 5,233.00 | 1.030 | 5 | I | R | F | P |
| I21 | 355.00 | 441.00 | 13.000 | 45 | I | A | R | F |
| I22 | 5,916.00 | 1,250.00 | 0.080 | 3 | I | R | F | F |
| I23 | 12,210.00 | 10,123.00 | 0.116 | 14 | I | R | U | N |
| I24 | 2,040.00 | 1,521.00 | 5.800 | 23 | I | R | F | F |
| I25 | 1,529.00 | 2,040.00 | 0.500 | 1 | I | R | F | F |
| I26 | 5,639.00 | 16,218.00 | 0.472 | 20 | I | R | F | N |
| I27 | 4,488.00 | 6,712.00 | 0.380 | 19 | I | V | F | N |
| I28 | 5,330.00 | 7,099.00 | 0.790 | 17 | I | R | F | P |
| I29 | 3,784.00 | 51,000.00 | 0.130 | 5 | P | V | U | N |
| I30 | 5,450.00 | 17,395.00 | 0.620 | 4 | I | R | F | P |

# TABLE4: FLORIDA DATA

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F1 | 69,421.00 | 102,064.00 | 0.0872 | 5 | I | C | U | F |
| F2 | 29,912.00 | 39,912.00 | 0.1596 | 41 | I | R | U | F |
| F3 | 23,743.00 | 26,520.00 | 0.1250 | 3 | I | V | U | F |
| F4 | 21,683.00 | 762.00 | 0.0963 | 22 | I | R | U | F |
| F5 | 24,912.00 | 70,522.00 | 0.0709 | 12 | I | R | U | F |
| F6 | 43,592.00 | 32,051.00 | 0.0780 | 1 | I | C | U | F |
| F7 | 30,695.00 | 103,736.00 | 0.0803 | 10 | I | R | U | F |
| F8 | 90,552.00 | 82,410.00 | 0.1083 | 8 | I | R | U | F |
| F9 | 33,470.00 | 104,522.00 | 0.1172 | 35 | I | R | U | F |
| F10 | 21,785.00 | 40,388.00 | 0.1238 | 14 | I | R | U | F |
| F11 | 18,992.00 | 34,149.00 | 0.2798 | 14 | I | V | U | F |
| F12 | 21,341.00 | 56,241.00 | 0.0697 | 30 | I | R | U | F |
| F13 | 117,926.00 | 80,737.00 | 0.3716 | 10 | I | C | U | F |
| F14 | 27,852.00 | 71,849.00 | 0.1590 | 4 | I | R | U | F |
| F15 | 27,931.00 | 70,037.00 | 0.1903 | 24 | I | R | U | F |
| F16 | 27,832.00 | 46,500.00 | 0.3372 | 25 | I | R | U | F |
| F17 | 30,075.00 | 75,658.00 | 0.1101 | 17 | I | R | U | F |

# TABLE 4: FLORIDA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F18 | 35,162.00 | 54,668.00 | 0.1628 | 9 | I | C | U | F |
| F19 | 4,775.00 | 6,569.00 | 0.3879 | 32 | I | R | U | F |
| F20 | 6,182.00 | 11,751.00 | 0.2511 | 35 | I | V | U | F |
| F21 | 1,571.00 | 3.030.00 | 1.9000 | 22 | I | R | U | F |
| F22 | 2,741.00 | 2,864.00 | 3 6700 | 17 | I | R | U | F |
| F23 | 1,950.00 | 14,786.00 | 0.3314 | 20 | I | C | U | F |
| F24 | 27,002.00 | 63,226.00 | 0.1085 | 11 | I | R | U | F |
| F25 | 8,896.00 | 16,117.00 | 0.3101 | 22 | I | R | U | F |
| F26 | 5,512.00 | 4,260.00 | 0.1878 | 12 | I | R | U | F |
| F27 | 5,537.00 | 5,320.00 | 0.0921 | 27 | I | R | U | F |
| F28 | 5,654.00 | 5,594.00 | 0.0876 | 36 | I | V | U | F |
| F29 | 5,720.00 | 11,194.00 | 0.0879 | 39 | I | R | U | F |
| F30 | 5,510.00 | 8,626.00 | 0.0946 | 1 | I | C | U | F |
| F31 | 5,510.00 | 4,387.00 | 0.0930 | 4 | I | R | U | F |
| F32 | 5,512.00 | 5,340.00 | 0.1910 | 2 | I | R | U | F |
| F33 | 9,321.00 | 38,754.00 | 0.1316 | 1 | I | R | U | F |
| F34 | 12,423.00 | 7,694.00 | 0.0464 | 1 | I | R | U | F |

# TABLE 4 : FLORIDA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F34 | 12,423.00 | 7,694.00 | 0.0464 | 1 | I | R | U | F |
| F35 | 9,692.00 | 15,624.00 | 0.3917 | 1 | I | R | U | F |
| F36 | 19,062.00 | 13,565.00 | 0.0650 | 26 | I | C | U | F |
| F37 | 5,932.00 | 2,990.00 | 0.1672 | 13 | I | R | U | F |
| F38 | 6,089.00 | 34,437.00 | 0.1992 | 36 | I | R | U | F |
| F39 | 8,892.00 | 20,559.00 | 0.1001 | 25 | I | R | U | F |
| F40 | 21,540.00 | 26,674.00 | 2.7180 | 12 | I | V | U | F |
| F41 | 5,220.00 | 1,825.00 | 0.0537 | 4 | I | R | U | F |
| F42 | 8,327.00 | 6,243.00 | 0.0471 | 1 | I | R | U | F |
| F43 | 67,484.00 | 51,836.00 | 0.0757 | 11 | I | C | U | F |
| F44 | 13,343.00 | 39,507.00 | 0.0894 | 45 | I | R | U | F |
| F45 | 4,692.00 | 16,780.00 | 0.2460 | 24 | I | R | U | F |
| F46 | 13,600.00 | 16,810.00 | 0.0210 | 24 | I | C | U | F |
| F47 | 17,556.00 | 68,410.00 | 0.1491 | 7 | I | R | U | F |
| F48 | 24,022.00 | 14,076.00 | 0.0342 | 1 | I | C | U | F |
| F49 | 24,022.00 | 13,076.00 | 0.0390 | 1 | I | C | U | F |
| F50 | 36,021.00 | 42,006.00 | 0.2064 | 9 | I | C | U | F |

# TABLE 4: FLORIDA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F51 | 114.00 | 134.00 | 109.8700 | 11 | I | A | R | F |
| F52 | 162.00 | 188.00 | 406.2500 | 4 | I | V | R | F |
| F53 | 222.00 | 98.00 | 163.1100 | 36 | I | V | R | F |
| F54 | 499.00 | 6,928.00 | 7.0000 | 32 | I | A | R | F |
| F55 | 276.00 | 9,538.00 | 18.0000 | 29 | I | A | R | F |
| F56 | 227.00 | 151.00 | 19.9000 | 4 | I | A | R | F |
| F57 | 143.00 | 5,497.00 | 3.0000 | 36 | I | A | R | F |
| F58 | 150.00 | 58.00 | 33.6600 | 19 | I | A | R | F |
| F59 | 147.00 | 2,135.00 | 36.8000 | 31 | I | A | R | F |
| F60 | 3,064.00 | 7,398.00 | 6.9800 | 32 | P | A | R | F |
| F61 | 1,909.00 | 5,924.00 | 0.9925 | 29 | P | V | R | F |
| F62 | 483.00 | 754.00 | 1.3000 | 27 | I | R | R | F |
| F63 | 357.00 | 588.00 | 3.0000 | 36 | I | R | R | F |
| F64 | 821.00 | 1,275.00 | 1.6000 | 6 | I | R | R | F |
| F65 | 339.00 | 2,918.00 | 18.0000 | 41 | I | V | R | F |
| F66 | 525.00 | 3,876.00 | 3.5900 | 25 | I | R | R | F |

# TABLE 4: FLORIDA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F67 | 461.00 | 3,385.00 | 3.5900 | 25 | I | R | R | F |
| F68 | 1,044.00 | 1,038.00 | 0.8670 | 12 | I | V | R | F |
| F69 | 1,020.00 | 87,830.00 | 0.1060 | 41 | I | R | R | F |
| F70 | 1,465.00 | 9,074.00 | 0.5400 | 40 | I | V | R | F |
| F71 | 1,323.00 | 5,528.00 | 9.9500 | 13 | I | V | R | F |
| F72 | 671.00 | 425.00 | 8.6320 | 7 | I | V | R | F |
| F73 | 714.00 | 1,102.00 | 4.0000 | 43 | I | R | R | F |
| F74 | 459.00 | 712.00 | 8.2640 | 41 | I | V | R | F |
| F75 | 1,062.00 | 717.00 | 0.4100 | 40 | I | V | R | F |
| F76 | 337.00 | 580.00 | 8.7870 | 4 | I | V | R | F |
| F77 | 231.00 | 238.00 | 4.2780 | 3 | I | V | R | F |
| F78 | 205.00 | 978.00 | 8.5130 | 35 | I | V | R | F |
| F79 | 6,142.00 | 446.00 | 0.5720 | 3 | I | R | R | F |
| F80 | 1,292.00 | 3,454.00 | 0.1100 | 67 | I | R | R | F |
| F81 | 643.00 | 303.00 | 2.8010 | 30 | I | R | R | F |
| F82 | 927.00 | 2,222.00 | 0.0900 | 28 | I | R | R | F |
| F83 | 125.00 | 98.00 | 0.9950 | 31 | I | A | R | F |

# TABLE4: FLORIDA DATA (Continued)

| PARCEL ID. | UNIT LAND VALUES IN CONSTANT DOLLARS PER ACRE | | SIZE OF PARCELS (ACRES) | TIME ELAPSED (MONTHS) | TYPE OF HIGHWAY | TYPE OF LAND USE | TYPE OF AREA | TYPE OF ACCESS CONTROL |
|---|---|---|---|---|---|---|---|---|
| | BEFORE | AFTER | | | | | | |
| F84 | 318.00 | 887.00 | 1.1500 | 23 | I | V | R | F |
| F85 | 421.00 | 1,326.00 | 1.5000 | 25 | I | R | R | F |
| F86 | 442.00 | 638.00 | 4.0000 | 11 | I | R | R | F |
| F87 | 1,742.00 | 1,176.00 | 5.0000 | 25 | I | A | R | F |
| F88 | 408.00 | 408.00 | 9.3090 | 2 | I | V | R | F |
| F89 | 2,386.00 | 2,278.00 | 8.3900 | 28 | I | A | R | F |
| F90 | 3,214.00 | 1,506.00 | 1.6600 | 8 | I | A | R | F |
| F91 | 3,406.00 | 389.00 | 3.7800 | 25 | I | R | R | F |
| F92 | 3,732.00 | 6,841.00 | 0.8200 | 2 | I | A | R | F |
| F93 | 1,572.00 | 1,185.00 | 0.2480 | 25 | I | R | R | F |
| F94 | 735.00 | 1,176.00 | 1.0000 | 16 | I | V | R | F |
| F95 | 170.00 | 176.00 | 221.8420 | 37 | I | A | R | F |
| F96 | 308.00 | 353.00 | 18.0000 | 44 | I | V | R | F |
| F97 | 311.00 | 362.00 | 42.7640 | 7 | I | V | R | F |
| F98 | 309.00 | 343.00 | 33.6060 | 53 | I | V | R | F |
| F99 | 516.00 | 129.00 | 0.7750 | 22 | I | V | R | F |
| F100 | 259.00 | 428.00 | 34.4000 | 52 | I | V | R | F |

# CHAPTER SIX

# RESULTS OF MODEL DEVELOPMENT

The results obtained in this research are presented by a series of mathematical models describing the relationship between a measure of highway economic impacts on land value and a set of explanatory variables. These results pertain to the analysis of land value data from the states of Indiana and Florida.

## Indiana (I) Set of Models

### Model I-CCl

The various statistics of this model are given in Table 5 . The multiple correlation coefficient was 0.9242. Variables associated with highway type, size, type of land use, time, type of area, and type of access control accounted for 85.42% of the variation in the dependent variable. The most significant of these variables, as shown by their respective beta coefficients is the variable denoting the type of highway improvement. Variables representing different types of highway improvements entered at the second and third steps. Their joint contribution to the statistical strength of the model as measured by the increase in the coefficient of multiple determination was appreciable.

Table 5 shows that after the seventh variable was entered into the model there was little increase in the coefficient of multiple determination. However, at the ninth step the lowest standard error of estimate was achieved.

# TABLE 5 : SUMMARY OF MODEL I-CCI

Dependent Variable Name : DDIFF

Mean : 92.79233

Standard Deviation : 214.6948

C.V. (at last step) : 112.13%

C.V. (lowest) : 108.71%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 268.46439 | 80.99414 | 0.5375 | 148.5875 |
| 2 | INTERS | -408.21287 | 89.09055 | 0.6576 | 130.1927 |
| 3 | PRIM | -383.80284 | 107.71778 | 0.7134 | 121.3938 |
| 4 | COMM | 168.09115 | 138.74092 | 0.7395 | 118.014 |
| 5 | PART | 196.19128 | 66.84969 | 0.7716 | 112.7895 |
| 6 | RES | -219.36664 | 97.40547 | 0.7964 | 108.7662 |
| 7 | RURAL | -109.44947 | 63.28418 | 0.8288 | 102.0001 |
| 8 | FULL | 115.18695 | 73.79647 | 0.8394 | 101.1038 |
| 9 | AGRIC | -148.08407 | 120.45654 | 0.8477 | 100.876 |
| 10 | TIME | 133.86304 | 156.34083 | 0.8523 | 101.929 |
| 11 | SIZE | 0.26248 | 0.54329 | 0.8542 | 104.0497 |

Model I-CC2

This model is summarized in Table 6 . The only difference between the variables in this model and those in the preceding one is that the logarithm transformations of the quantitative independent variables in model I-CC1 were used. The order in which the independent variables entered the model was an exact replica of what occurred in model I-CC1. However, there were slight improvements both in $R^2$ (86.31% for I-CC2 as against 85.42% for I-CC1) and in the coefficient of variation (coefficient of variation at last step is 108.65% here as against 112.13% for model I-CC1).

Model I-CC3

Table 7 summarizes the results of this model. The dependent variable is the same as in the other two models. The difference here is in the treatment of the independent variables. The quantitative independent variables (size and time) were not assumed to be linearly related to the dependent variable. Instead the range of each variable was broken down into dummy classes to trace any curvilinearity that might be present. This model from a statistical viewpoint is better than all the others developed using DDIFF as the dependent variable. It produced the highest value of the coefficient of multiple correlation (R=0.9341) and the lowest coefficient of variation at the last step (C.V=101.95%). However, it included two variables more than the other two.

The dominant factors were in their order of importance, type of area, highway type, and type of land use.

# TABLE 6 : SUMMARY OF MODEL 1-CC2

Dependent Variable Name : DDIFF

Mean : 92.79233

Standard Deviation : 214.6948

C.V. (at last step) : 108.6534%

C.V. (lowest) : 106.9223%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 269.15546 | 79.62573 | 0.5375 | 148.5875 |
| 2 | INTERS | -444.65859 | 88.30578 | 0.6576 | 130.1927 |
| 3 | PRIM | -404.24671 | 98.73496 | 0.7134 | 121.3938 |
| 4 | COMM | 147.49698 | 136.05028 | 0.7395 | 118.0141 |
| 5 | PART | 186.25349 | 66.72860 | 0.7716 | 112.7895 |
| 6 | RES | -219.74781 | 94.33022 | 0.7964 | 108.7662 |
| 7 | RURAL | -120.03859 | 72.67954 | 0.8288 | 102.0001 |
| 8 | FULL | 126.03911 | 72.74212 | 0.8394 | 101.1038 |
| 9 | AGRIC | -182.62240 | 121.89215 | 0.8477 | 100.8760 |
| 10 | TIME | 55.13631 | 47.82991 | 0.8601 | 99.2157 |
| 11 | SIZE | 28.76026 | 45.50754 | 0.8631 | 100.8220 |

# TABLE 7 : SUMMARY OF MODEL 1-CC3

Dependent Variable Name : DD1FF

Mean : 92.79233

Standard Deviation : 214.6948

C.V. (at last step) : 111.2%

C.V. (lowest) : 101.947%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 293.11379 | 79.30863 | 0.5375 | 148.5875 |
| 2 | INTERS | -395.16299 | 91.29794 | 0.6576 | 130.1927 |
| 3 | TIME B | 61.81690 | 72.97407 | 0.7142 | 121.2244 |
| 4 | PRIM | -353.19096 | 107.89773 | 0.7521 | 115.1197 |
| 5 | COMM | 212.69522 | 139.83425 | 0.7781 | 111.1747 |
| 6 | PART | 188.49165 | 66.60544 | 0.8144 | 103.8677 |
| 7 | RES | -166.87543 | 100.96005 | 0.8332 | 100.658 |
| 8 | RURAL | -104.89063 | 69.58132 | 0.8534 | 96.5906 |
| 9 | FULL | 102.07509 | 73.30343 | 0.8661 | 94.5996 |
| 10 | AGRIC | -91.00812 | 120.01364 | 0.8697 | 95.7318 |
| 11 | TIME A | -30.83178 | 68.30307 | 0.8708 | 97.9494 |
| 12 | SIZE A | -59.21092 | 126.76886 | 0.8716 | 100.4901 |
| 13 | SIZE B | -40.74515 | 116.46415 | 0.8725 | 103.1889 |

Model I-PC1

The results obtained with this model are summarized in Table 8. All the independent variables considered were entered into the model. The variable denoting one of the classes of highway type was the first to enter the model and it alone explained about 40% of the variation in the dependent variable. This model yielded a coefficient of multiple correlation of 0.9107. However at the seventh step in the development of this model, a coefficient of multiple correlation of 0.9044 was achieved. This means that the four additional variables added from step seven to step eleven increased the total variation explained by the model by only about one percent.

Model I-PC2

This model is presented in Table 9 . Of the eleven independent variables that were candidates for inclusion in the model, ten made it. Together these ten variables explained about 83.3% of the variation in the dependent variable. The first five variables to enter this model explained about 77% of this variation. This shows that the last five variables only increased the magnitude of $R^2$ by about six percent. This model uses less variables than model I-PC1 and appears to perform better from statistical considerations. It has a coefficient of variation of 114.4% as compared to 125.0% for model I-PC1 at the last step.

# TABLE 8 : SUMMARY OF MODEL I-PC I

Dependent Variable Name : PD I FF

Mean : 531.9

Standard Deviation : 1267.91242

C.V. (at last step) : 125.009 %

C.V. ( lowest ) : 116.523 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2818.31363 | 569.32772 | 0.4016 | 998.1304 |
| 2 | PRIM | -2339.42927 | 688.36393 | 0.5538 | 877.6996 |
| 3 | URBAN | 648.10184 | 517.58814 | 0.6019 | 844.8508 |
| 4 | PART | 837.44223 | 427.19886 | 0.6473 | 810.98 |
| 5 | RES | -365.69708 | 622.46371 | 0.6983 | 765.5050 |
| 6 | RURAL | -1376.74285 | 404.41369 | 0.7767 | 672.8008 |
| 7 | AGRIC | 1177.28358 | 769.77021 | 0.8180 | 621.0215 |
| 8 | TIME | 957.93216 | 999.08655 | 0.8270 | 619.7898 |
| 9 | COMM | 273.90543 | 886.61542 | 0.8281 | 633.0762 |
| 10 | SIZE | 1.07418 | 3.47186 | 0.8289 | 647.8541 |
| 11 | FULL | -90.76121 | 471.59186 | 0.8293 | 664.9231 |

# TABLE 9 : SUMMARY OF MODEL 1-PC2

Dependent Variable Name : PDIFF

Mean : 531.9

Standard Deviation: 1267.91242

C.V.(at last step): 120.5134%

C.V.(lowest): 114.3599%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2994.72462 | 547.21031 | 0.4016 | 998.1304 |
| 2 | PRIM | -2475.04426 | 624.72295 | 0.5538 | 877.6996 |
| 3 | TIME | 344.11872 | 297.90194 | 0.6337 | 810.4500 |
| 4 | RURAL | -1343.01438 | 461.01796 | 0.6910 | 759.0392 |
| 5 | AGRIC | 1080.44820 | 660.91789 | 0.7669 | 672.8947 |
| 6 | PART | 831.48226 | 342.66106 | 0.8096 | 621.1804 |
| 7 | URBAN | 636.50996 | 504.78098 | 0.8254 | 608.2806 |
| 8 | RES | -367.81995 | 565.01660 | 0.8317 | 611.2749 |
| 9 | COMM | 222.44746 | 864.98688 | 0.8324 | 625.0807 |
| 10 | SIZE | 39.02970 | 288.70333 | 0.8325 | 641.0111 |

Model I-PC3

A summary of the results obtained in the development of this model is shown in Table 10 . At the last step, a coefficient of multiple correlation (r) of 0.9111 was obtained. The table also reveals that the lowest value of the standard error of estimate was obtained at the seventh step. The increase in the coefficient of multiple determination from the seventh step to the last (thirteenth) step was only about one percent. This indicates that we can explain about 82% of the variations in the response variable by considering only seven variables. Of this 82%, the variables denoting the type of highway improvement accounted for about 55%.

Model I-RR1

Table 11 summarizes the results obtained in the development of this model. An inspection of Table 11 reveals that the standard error continued to decrease from its value at the first step until the eighth step was reached. Then it began to increase again. Also at this eighth step a value of 0.8270 was obtained for the coefficient of multiple determination. The increase in $R^2$ from this stage to the last stage is negligible (less than a quarter of a percent). Thus the inclusion of variables into the final model could be stopped at this stage without any significant loss of information on the observed variation in the response variable.

# TABLE 10 : SUMMARY OF MODEL 1-PC3

Dependent Variable Name : PDIFF

Mean : 531.9

Standard Deviation : 1267.91242

C.V. (at last step) : 132.28%

C.V. (lowest) : 116.755 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2773.29484 | 622.53331 | 0.4016 | 998.1304 |
| 2 | PRIM | -2322.50449 | 735.72235 | 0.5538 | 877.6996 |
| 3 | URBAN | 822.54695 | 540.78181 | 0.6019 | 844.8508 |
| 4 | PART | 815.18203 | 454.16258 | 0.6473 | 810.9800 |
| 5 | RES | -180.79376 | 688.41640 | 0.6983 | 765.505 |
| 6 | RURAL | -1319.52839 | 474.4542 | 0.7767 | 672.8008 |
| 7 | AGRIC | 1499.6598 | 818.33707 | 0.8180 | 621.0215 |
| 8 | TIME A | -207.50434 | 465.73820 | 0.825 | 623.3379 |
| 9 | COMM | 515.62746 | 953.48788 | 0.8276 | 633.9692 |
| 10 | TIME B | 149.29397 | 497.58836 | 0.8285 | 648.7831 |
| 11 | FULL | -143.05942 | 499.83412 | 0.8294 | 664.7512 |
| 12 | SIZE B | -183.55088 | 749.13420 | 0.8300 | 682.8463 |
| 13 | SIZE A | -91.74582 | 864.39894 | 0.8301 | 703.6142 |

# TABLE 11 : SUMMARY OF MODEL I-RR 1

Dependent Variable Name : R VALUE
Mean : 631.85667
Standard Deviation : 1267.9632
C.V. (at last step ) : 105.24%
C.V. ( lowest ) : 98.096 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2818.27098 | 569.36132 | 0.4017 | 998.1485 |
| 2 | PRIM | -2339.26762 | 688.40456 | 0.5538 | 877.7554 |
| 3 | URBAN | 648.07755 | 517.61869 | 0.6019 | 844.9082 |
| 4 | PART | 837.38632 | 427.22407 | 0.6473 | 811.0365 |
| 5 | RES | -305.75083 | 622.50045 | 0.6983 | 765.5354 |
| 6 | RURAL | -1376.77998 | 404.43757 | 0.7767 | 672.8361 |
| 7 | AGRIC | 1177.52089 | 769.81565 | 0.8180 | 621.0580 |
| 8 | TIME | 957.88982 | 999.14552 | 0.8270 | 619.8303 |
| 9 | COMM | 273.73184 | 886.66776 | 0.8281 | 633.1198 |
| 10 | SIZE | 1.07491 | 3.47207 | 0.8289 | 647.8969 |
| 11 | FULL | -91.07414 | 471.61969 | 0.8293 | 664.9623 |

Model I-RR2

The results obtained in developing this model are presented in Table 12 . The ten independent variables included in the final model explained about 83.3% of the variation in the response variable. There was no significant increase in $R^2$ from the seventh step to the last step. Therefore the variables entering after the seventh stage contributed little information on the observed variations.

Computation of the beta coefficients showed that the variable denoting highway type has the strongest influence on the values of the dependent variable.

Of all the models using recovery rate as their measure of highway economic impact, this model appears to be the best. It produces the highest $R^2$ and next to model I-RR2A is the one having the lowest coefficient of variation.

Model I-RR3

Table 13 gives a summary of the results obtained in the development of this model. Of all the models using recovery rate as their measure of impact, this model included the highest number of variables. Yet it has an $R^2$ of 0.8301 which is slightly below that of model I-RR2 ($R^2$ = 0.8325) which includes only ten variables at the final stage. Table 13 shows that any attempt to decrease the standard error of estimate beyond stage seventh would be futile. Also beyond this stage, any variable added to the model was of no advantage in explaining the variation in the recovery rate.

# TABLE 12 : SUMMARY OF MODEL I-RR2

Dependent Variable Name : R VALUE

Mean : 631.85667

Standard Deviation : 1267.9632

C.V. (at last step) : 101.4539 %

C.V. ( lowest ) : 96.27449 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2994.83294 | 547.23783 | 0.4017 | 998.1485 |
| 2 | PRIM | -2474.93902 | 624.75437 | 0.5538 | 877.7554 |
| 3 | TIME | 344.16820 | 297.91693 | 0.6337 | 810.4883 |
| 4 | RURAL | -1343.12641 | 461.04114 | 0.6910 | 759.0938 |
| 5 | AGRIC | 1080.28680 | 660.95113 | 0.7669 | 672.9422 |
| 6 | PART | 831.55110 | 342.67830 | 0.8096 | 621.2177 |
| 7 | URBAN | 636.54093 | 504.80637 | 0.8254 | 608.3168 |
| 8 | RES | -368.02012 | 565.04502 | 0.8317 | 611.3055 |
| 9 | COMM | 222.20557 | 865.03039 | 0.8324 | 625.1141 |
| 10 | SIZE | 39.15682 | 288.71785 | 0.8325 | 641.0433 |

# TABLE 13 : SUMMARY OF MODEL 1-RR3

Dependent Variable Name : R VALUE

Mean : 631.88

Standard Deviation : 1267.96745

C.V. (at last step) : 111.35 %

C.V. (lowest) : 98.283 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2773.57084 | 622.54421 | 0.4017 | 998.1239 |
| 2 | PRIM | -2322.62657 | 735.73523 | 0.5539 | 877.6892 |
| 3 | URBAN | 822.34074 | 540.79129 | 0.602 | 844.8571 |
| 4 | PART | 814.99156 | 454.17053 | 0.6473 | 810.9950 |
| 5 | RES | -180.83059 | 68842846 | 0.6984 | 765.4914 |
| 6 | RURAL | -1319.43330 | 474.46251 | 0.7767 | 672.8005 |
| 7 | AGRIC | 1499.88486 | 818.35140 | 0.818 | 621.0313 |
| 8 | TIME A | -207.31273 | 465.74636 | 0.825 | 623.3559 |
| 9 | COMM | 515.34394 | 953.50458 | 0.8276 | 633.9943 |
| 10 | TIME B | 149.47869 | 497.59707 | 0.8285 | 648.8049 |
| 11 | FULL | -143.47094 | 499.84288 | 0.8294 | 664.7636 |
| 12 | SIZE B | -183.61805 | 794.14811 | 0.8300 | 682.8586 |
| 13 | SIZE A | -91.80759 | 864.41408 | 0.8301 | 703.6265 |

### Model I-RR2A

The independent variables considered for inclusion into this model are exactly the same as those that were used for model I-RR2. The only difference in the two models is that the dependent variable here is the logarithm transformation of the dependent variable in model I-RR2. Only eight of the initial list of independent variables entered this model. Together they accounted for about 78% of the variation in the dependent variable. This is lower than the value of $R^2$ obtained at a corresponding stage in model I-RR2 ($R^2 = 0.8317$). This model however has a coefficient of variation of 16.5% at the last step as compared to 101.5% for model I-RR2, (See Figure 14).

## Indiana-Florida (IF) Set of Models

### Model IF-CC1

The results obtained at the various stages in the development of this model are shown in Table 15 . The model explained only about 31% of the variation in the dependent variable. A high coefficient of variation of about 203.6% was obtained at the last stage.

### Model IF-CC2

This model was summarized in Table 16 shows that the logarithm transformation of two of the independent variables in model IF-CC1 helped only in increasing the value of $R^2$ by about four percent. It also provided a decrease of about the same amount in the coefficient of variation.

# TABLE 14 : SUMMARY OF MODEL 1-RR2A

Dependent Variable Name : R VALUE

Mean : 2.2272

Standard Deviation : 0.66613

C.V. ( at last step ) : 16.5 %

C.V. ( lowest ) : 16.1 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -1.57635 | 0.28521 | 0.4482 | 0.5036 |
| 2 | RURAL | -0.94238 | 0.25379 | 0.539 | 0.4687 |
| 3 | PRIM | -1.19708 | 0.32447 | 0.6422 | 0.4208 |
| 4 | RES | -0.28447 | 0.31283 | 0.7208 | 0.3791 |
| 5 | FULL | -0.49475 | 0.25663 | 0.7481 | 0.3675 |
| 6 | SIZE | 0.19250 | 0.15835 | 0.7699 | 0.3588 |
| 7 | AGRIC | 0.32686 | 0.39631 | 0.7733 | 0.3641 |
| 8 | PART | -0.18365 | 0.23789 | 0.7796 | 0.3675 |

# TABLE 15 : SUMMARY OF MODEL IF-CCI

Dependent Variable Name : DDIFF

Mean : 80.76685

Standard Deviation : 190.46558

C.V. (at last step) : 203.63%

C.V. ( lowest ) : 201.579%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 140.03458 | 64.45359 | 0.1039 | 180.9981 |
| 2 | COMM | -148.64252 | 66.00867 | 0.1686 | 175.0339 |
| 3 | FULL | -42.42460 | 79.22044 | 0.2242 | 169.7415 |
| 4 | TIME | 186.90442 | 103.65200 | 0.2582 | 166.6532 |
| 5 | INTERS | -327.67912 | 120.04271 | 0.2723 | 165.7217 |
| 6 | PRIM | -282.15205 | 124.77518 | 0.3033 | 162.8093 |
| 7 | PART | 39.36177 | 85.09568 | 0.3050 | 163.2782 |
| 8 | RES | 14.81068 | 38.60035 | 0.3062 | 163.8089 |
| 9 | RURAL | -12.66481 | 68.30545 | 0.3064 | 164.4665 |

# TABLE 16 : SUMMARY OF MODEL IF-CC2

Dependent Variable Name : DD1FF

Mean : 80.76685

Standard Deviation : 190.46558

C.V. (at last step) : 199.563%

C.V. ( lowest ) : 197.6958%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 106.66534 | 43.13643 | 0.1039 | 180.9981 |
| 2 | COMM | -175.26933 | 69.57493 | 0.1686 | 175.0339 |
| 3 | FULL | -25.84728 | 69.28991 | 0.2242 | 169.7415 |
| 4 | TIME | 78.04616 | 30.24173 | 0.2610 | 166.3321 |
| 5 | SIZE | -46.25273 | 24.66407 | 0.2766 | 165.2274 |
| 6 | INTERS | -356.37521 | 108.55894 | 0.2952 | 163.7525 |
| 7 | PRIM | -312.17827 | 112.54208 | 0.3353 | 159.6727 |
| 8 | PART | 57.43989 | 83.88241 | 0.3377 | 160.0415 |
| 9 | AGRIC | 17.61217 | 46.50503 | 0.3391 | 160.5442 |
| 10 | RES | -10.17760 | 43.89916 | 0.3394 | 161.1810 |

### Model IF-CC3

The various statistics of this model are given in Table 17 . The measure of correlation was expressed by a multiple correlation coefficient of 54.66%. The variables associated with time after acquisition, type of highway improvement, type of area, type of land use and type of access control accounted for only about 30% of the variation in the dependent variable. The fact that the two quantitative independent variables (size, and time) were broken down into dummy classes did not seem to be helpful in improving the statistical showing of this model.

### Model IF-PC1

This model is summarized in Table 18 . All the eleven independent variables that were considered passed the test for inclusion into the model. Together they explained about 46% of the variation in the dependent variable. Judged by the amount of variation in the dependent variable explained by the models, this model ranks second among the IF models. However, it has a high value of 221.5% for its percent standard error.

### Model IF-PC2

Table 19 provides a summary of the results obtained in the development of this model. Of all the models developed under the IF series, this model gave the highest value of the coefficient of multiple determination. The ten independent variables in the final form of this model accounted for 46.40% of the variation in the dependent variable. The precision of this model as indicated by the percent standard error was about 219.2%.

# TABLE 17 : SUMMARY OF MODEL IF-CC3

Dependent Variable Name : DDIFF

Mean : 80.76685

Standard Deviation : 190.46558

C.V. (at last step) : 204.75479%

C.V. (lowest) : 202.674%

| STEP NO | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | URBAN | 148.18753 | 37.69003 | 0.1039 | 180.998 |
| 2 | COMM | -155.38454 | 65.01236 | 0.1686 | 175.0339 |
| 3 | FULL | -48.39383 | 70.18222 | 0.2242 | 169.7415 |
| 4 | TIME A | -23.23395 | 58.85415 | 0.2431 | 168.333 |
| 5 | INTERS | -346.47984 | 113.03744 | 0.2584 | 167.2933 |
| 6 | PRIM | -303.38319 | 117.80473 | 0.2957 | 163.6938 |
| 7 | PART | 36.66317 | 85.55218 | 0.2972 | 164.1942 |
| 8 | TIME B | 20.82775 | 58.07887 | 0.2980 | 164.7724 |
| 9 | RES | 12.97200 | 37.23823 | 0.2987 | 165.3740 |

# TABLE 18 : SUMMARY OF MODEL IF-PC I

Dependent Variable Name : PD I FF

Mean : 271.52

Standard Deviation : 780.81635

C.V. ( at last step ) : 221.475 %

C.V. ( lowest ) : 219.658 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2516.12954 | 440.01933 | 0.1716 | 713.4212 |
| 2 | PRIM | -2223.10674 | 456.84424 | 0.3246 | 646.7368 |
| 3 | AGRIC | 388.22338 | 170.73960 | 0.3674 | 628.3632 |
| 4 | PART | 712.56587 | 312.26346 | 0.3966 | 616.1632 |
| 5 | TIME | 1102.52330 | 380.57924 | 0.4309 | 600.7906 |
| 6 | RURAL | -169.12011 | 250.50968 | 0.4410 | 597.8650 |
| 7 | RES | -267.14409 | 160.13869 | 0.4482 | 597.4158 |
| 8 | SIZE | -1.16557 | 1.18149 | 0.4519 | 596.8934 |
| 9 | COMM | -254.60259 | 250.95325 | 0.4563 | 596.9501 |
| 10 | URBAN | 118.61808 | 237.14970 | 0.4572 | 598.9481 |
| 11 | FULL | -66.23809 | 292.28102 | 0.4574 | 601.3498 |

# TABLE 19 : SUMMARY OF MODEL IF-PC 2

Dependent Variable Name : PD I FF

Mean : 271.52

Standard Deviation : 780.81635

C.V. (at last step ) : 219.1956 %

C.V. ( lowest ) : 218.158 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2695.68251 | 362.83019 | 0.1716 | 713.4212 |
| 2 | PRIM | -2395.27692 | 408.10197 | 0.3246 | 646.7368 |
| 3 | AGRIC | 412.86883 | 170.73740 | 0.3674 | 628.3632 |
| 4 | TIME | 364.08498 | 112.83687 | 0.3992 | 614.8206 |
| 5 | PART | 755.98831 | 269.14656 | 0.4362 | 597.9828 |
| 6 | SIZE | -123.55409 | 92.53644 | 0.4471 | 594.5672 |
| 7 | RES | -304.09712 | 164.52176 | 0.4557 | 592.3637 |
| 8 | COMM | -298.12125 | 258.46893 | 0.4602 | 592.3442 |
| 9 | RURAL | -125.55588 | 221.53877 | 0.4639 | 592.7376 |
| 10 | URBAN | 33.32646 | 210.08250 | 0.4640 | 595.16 |

Model IF-PC3

A summary of the results obtained during the development of this model is shown in Table 20 . This model achieved a coefficient of multiple correlation of 0.6760 at the last step. Thus the thirteen independent variables in the final equation explained about 45.7% of the variation in the dependent variable. No substantial increase in the coefficient of multiple determination occurred after the eighth step.

Model IF-RR1

The results produced in developing this model are summarized in Table 21 . The eight variables in the final form of this model yielded a coefficient of multiple correlation of 0.4933. Table 21 shows that after the fifth independent variable had been entered into the model no substantial increase in the value of $R^2$ occurred. Also after the fifth stage no further decrease was observed in the value of the standard error estimate.

Model IF-RR2

Details of the results obtained in developing this model are presented in Table 22 . The logarithm transformations of the independent variables size and time were used in this model. This transformation increased the value of the coefficient of multiple determination obtained for model IF-RR1 by about three percent. About a two percent reduction in the percent standard error also resulted from this transformation.

# TABLE 20 : SUMMARY OF MODEL IF-PC3

Dependent Variable Name : PD I FF

Mean : 272.12077

Standard Deviation : 780.70208

C.V. (at last step) : 222.9588 %

C.V. (lowest) : 220.68 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2638.14611 | 445.24392 | 0.1715 | 713.387 |
| 2 | PRIM | -2366.24600 | 461.96527 | 0.3245 | 646.6988 |
| 3 | AGRIC | 383.86919 | 179.25390 | 0.3672 | 628.3999 |
| 4 | PART | 705.06509 | 315.76245 | 0.3963 | 616.2250 |
| 5 | TIME B | 207.82159 | 224.64376 | 0.4271 | 602.7145 |
| 6 | URBAN | 154.24258 | 238.87642 | 0.4333 | 601.8635 |
| 7 | RES | -293.25839 | 164.58802 | 0.4403 | 600.5794 |
| 8 | COMM | -296.8307 | 252.95993 | 0.4450 | 600.5350 |
| 9 | RURAL | -121.89741 | 252.05296 | 0.4476 | 601.5987 |
| 10 | TIME A | -120.07557 | 229.53773 | 0.4497 | 602.9639 |
| 11 | SIZE A | 351.36589 | 298.82758 | 0.4515 | 604.5601 |
| 12 | SIZE B | 329.08353 | 322.89842 | 0.4563 | 604.4327 |
| 13 | FULL | -102.50083 | 295.26796 | 0.4569 | 606.7174 |

# TABLE 21 :SUMMARY OF MODEL IF-RRI

Dependent Variable Name: R VALUE

Mean : 433.18769

Standard Deviation : 1062.85829

C.V. (at last step) : 220.363%

C.V. (lowest) : 218.337%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2374.45261 | 582.95800 | 0.0812 | 1022.7668 |
| 2 | PRIM | -2044.19002 | 647.72514 | 0.1637 | 979.5868 |
| 3 | TIME | 1611.28074 | 590.15183 | 0.2077 | 957.2619 |
| 4 | PART | 720.78257 | 387.61882 | 0.2285 | 948.4045 |
| 5 | AGRIC | 340.54010 | 239.12591 | 0.2388 | 945.8096 |
| 6 | SIZE | -1.32148 | 1.83466 | 0.2426 | 947.2971 |
| 7 | COMM | 96.28071 | 333.41412 | 0.2432 | 950.7505 |
| 8 | URBAN | 29.50866 | 200.84532 | 0.2434 | 954.5860 |

# TABLE 22 : SUMMARY OF MODEL IF-RR 2

Dependent Variable Name : R VALUE

Mean : 433.18769

Standard Deviation : 1062.85829

C.V. (at last step) : 218.126 %

C.V. (lowest) : 215.858 %

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2517.17474 | 576.00865 | 0.0812 | 1022.7668 |
| 2 | PRIM | -2189.73718 | 647.89382 | 0.1637 | 979.5868 |
| 3 | TIME | 493.20098 | 179.10769 | 0.2059 | 958.3632 |
| 4 | PART | 779.60298 | 427.30780 | 0.2237 | 951.3378 |
| 5 | AGRIC | 454.23609 | 271.04644 | 0.2339 | 948.8891 |
| 6 | SIZE | -354.55041 | 146.91535 | 0.2564 | 938.6351 |
| 7 | RURAL | 227.23443 | 351.76466 | 0.2680 | 935.0733 |
| 8 | RES | -151.08909 | 261.28202 | 0.2694 | 938.0096 |
| 9 | URBAN | -130.57555 | 333.64867 | 0.2704 | 941.2780 |
| 10 | COMM | -118.00155 | 410.27959 | 0.2709 | 944.8963 |

Model IF-RR3

The essential statistics of this model are reported in Table 23 . The model yielded a coefficient of multiple correlation of 0.4861 at the last step. No statistical improvement was achieved after the fifth variable was entered into the model.

Model IF-RR2A

A summary of the results obtained in developing this model is given in Table 24 . The ten independent variables that qualified for inclusion in this model explained about 24% of the variation in the dependent variable. The statistical importance of this model however lies in its low percent standard error of estimate which was about 22%.

# TABLE 23 : SUMMARY OF MODEL IF-RR3

Dependent Variable Name : R VALUE

Mean : 433.19308

Standard Deviation : 1062.86045

C.V. ( at last step ) : 222.305%

C.V. ( lowest ) : 219.444%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -2636.02501 | 587.67005 | 0.0812 | 1022.7642 |
| 2 | PRIM | -2348.65526 | 655.59931 | 0.1637 | 979.574 |
| 3 | TIME B | 346.74594 | 355.92646 | 0.1999 | 961.9485 |
| 4 | PART | 684.41509 | 387.97110 | 0.2180 | 954.7875 |
| 5 | AGRIC | 341.15742 | 269.89581 | 0.2311 | 950.6163 |
| 6 | SIZE A | 399.32808 | 466.90021 | 0.2325 | 953.5869 |
| 7 | SIZE B | 307.95398 | 511.92374 | 0.2354 | 955.6405 |
| 8 | TIME A | -98.34658 | 359.14451 | 0.2359 | 959.3084 |
| 9 | RES | -56.03969 | 210.24317 | 0.2363 | 963.0122 |

# TABLE 24 : SUMMARY OF MODEL IF-RR2A

Dependent Variable Name : R VALUE

Mean : 2.20713

Standard Deviation : 0.53670

C.V.(at last step): 22.06485%

C.V.(lowest): 21.79301%

| STEP NO. | INDEPENDENT VARIABLE | AT LAST STEP REPORTED | | $R^2$ AT EACH STEP SPECIFIED | STANDARD ERROR OF ESTIMATE |
|---|---|---|---|---|---|
| | | REGRESSION COEFFICIENT | STANDARD ERROR | | |
| 1 | INTERS | -1.09156 | 0.35243 | 0.0989 | 0.5115 |
| 2 | TIME | 0.30478 | 0.09240 | 0.1615 | 0.4953 |
| 3 | PRIM | -0.79500 | 0.36627 | 0.2072 | 0.4835 |
| 4 | SIZE | -0.10233 | 0.07435 | 0.2190 | 0.4819 |
| 5 | FULL | -0.17795 | 0.23524 | 0.2257 | 0.4817 |
| 6 | RES | -0.18405 | 0.12634 | 0.2342 | 0.4810 |
| 7 | COMM | -0.19115 | 0.20791 | 0.2389 | 0.4815 |
| 8 | RURAL | -0.08384 | 0.20489 | 0.2400 | 0.4831 |
| 9 | PART | -0.04902 | 0.25367 | 0.2402 | 0.4870 |
| 10 | URBAN | -0.03558 | 0.19606 | 0.2404 | 0.4870 |

# CHAPTER SEVEN

# DISCUSSION OF RESULTS

## Effect of Data Inadequacy on Results

The results obtained in this research are limited, as to general application, by the range of the data used in the calibration of the developed regression equations. There were only a few observations on some classes of the variables, thus restricting the interpretations and use of the models.

For the Florida data, as noted earlier, information was used from one hundred (100) parcels. Unfortunately all but four of these parcels were affected by interstate highway improvements, none were in the urban fringe area, and all were with full control of access.

For the Indiana data, thirty (30) parcels were included. Most of these were affected by an interstate highway improvement located in the urban fringe or rural areas and with a land use of residential or agricultural (see Figure 22 ). Each of the three types of access control was included by a reasonable sample but few parcel sales had occurred after an 18-month period and few parcels were larger than 15 acres (see Figure 23 ).

Although some classes of most variables thus contained only a few observations, it is felt the modeling effort of this research provides indications of important variables and thereby gives guidance and perhaps will foster the collection of sufficient data for eventual development of a reliable predictive model.

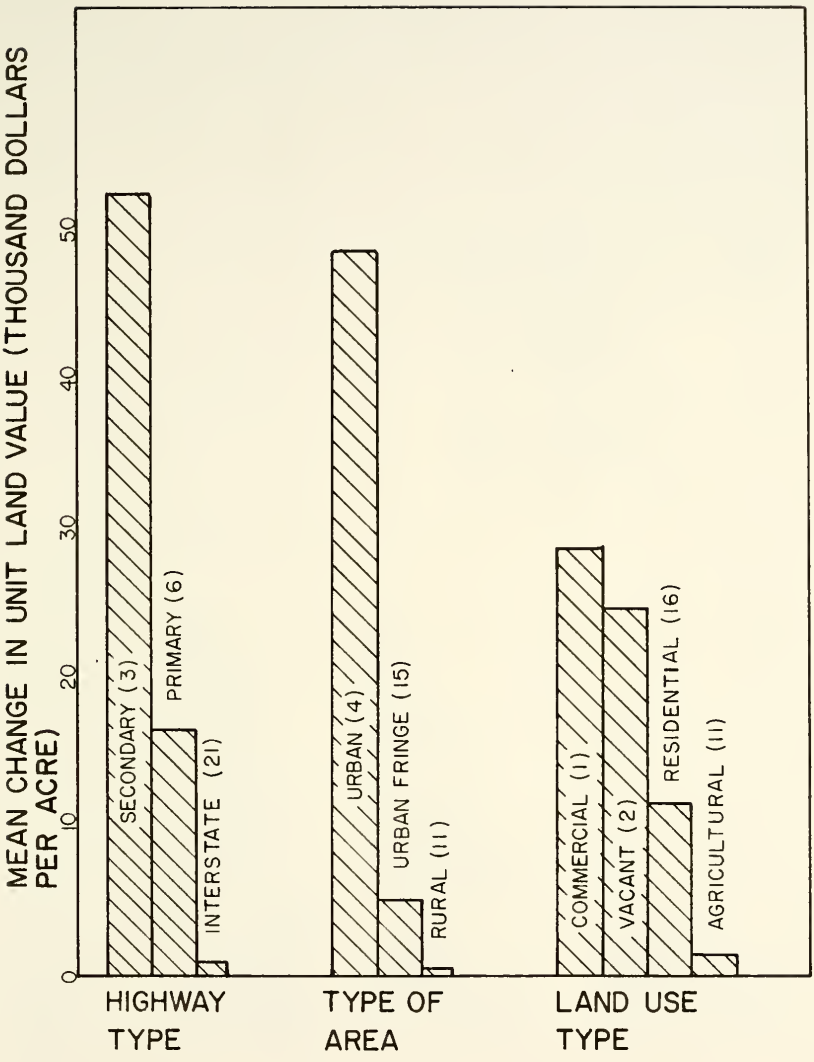



FIGURE 22 : MEAN CHANGE IN LAND VALUES FOR CLASSES WITHIN EACH VARIABLE

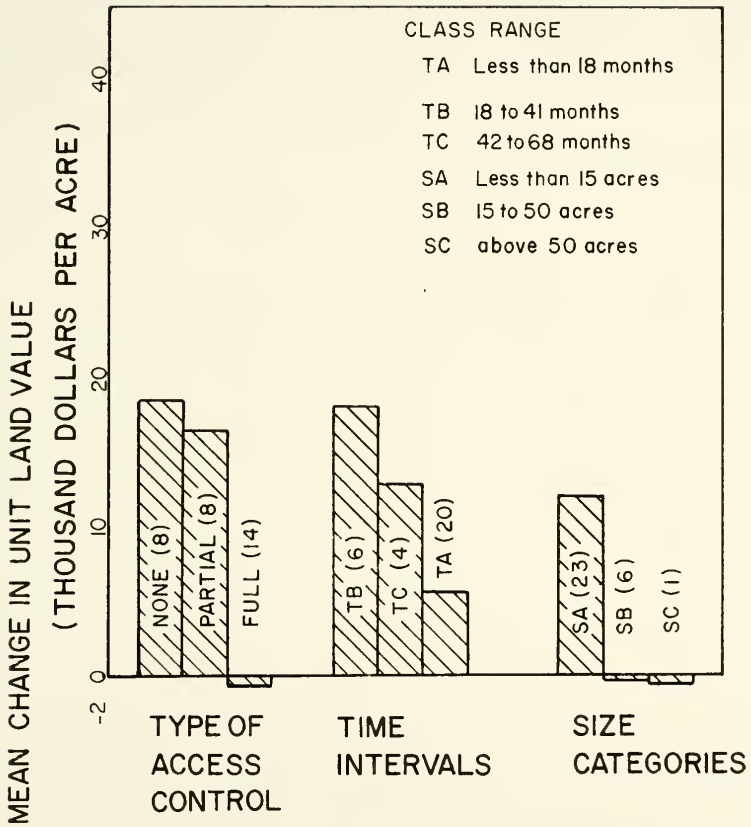



FIGURE 23: MEAN CHANGE IN LAND VALUES FOR CLASSES WITHIN EACH VARIABLE

The usefulness of the developed models is further restricted by the number of independent variables considered. For instance, distance to the highway improvement from the land parcel was not considered as a variable in this work because data were only available on parcels located adjacent to a highway improvement and sufficient time and effort were not available to collect original data. It is clear from previous research however, that this variable must be included in any comprehensive attempt at modeling the problem.

With these restrictions in mind the regression equations developed in this work will now be discussed.

## Comparison Between the I Set and the IF Set of Models

Models developed under the I set were statistically superior to those generated under the IF set. Since the same criteria were used in building these models, comparisons were limited to their statistical showings as reflected by their coefficients of multiple determination and their percent standard errors of estimate. Table 25 gives a comparative summary of the models developed in both sets.

Models in the IF set had $R^2$ values ranging from 0.2363 to 0.4640. In contrast, models in the I set had higher coefficients of multiple determination ranging from 0.7796 to 0.8725. The models in the IF set also exhibited a relatively higher standard error of estimate compared to models in the I set. Apart from model IF-RR2A, the models in the IF series had coefficients of variation ranging from 199.6% to 222.9%. The corresponding range for most of the models in the I set was from 101.5% to 132.3%. Model I-RR2A had the lowest coefficient of variation (C.V. = 16.5%) of all the models developed in this work.

| MODEL NAME | COEFFICIENT OF MULTIPLE DETERMINATION | COEFFICIENT OF VARIATION | NUMBER OF INDEPENDENT VARIABLES | $\left(\frac{F_o}{F_c}\right)$ |
|---|---|---|---|---|
| I - CC1 | 0.8542 | 112.13% | 11 | 4.6 |
| IF-CC1 | 0.3064 | 203.63% | 9 | — |
| I - CC2 | 0.8631 | 108.65% | 11 | 4.9 |
| IF-CC2 | 0.3394 | 199.56% | 10 | — |
| I - CC3 | 0.8725 | 111.20% | 13 | 4.0 |
| IF-CC3 | 0.2987 | 204.75% | 9 | — |
| I-PC1 | 0.8293 | 125.00% | 11 | 3.3 |
| IF-PC1 | 0.4574 | 221.48% | 11 | — |
| I - PC2 | 0.8325 | 120.51% | 10 | 4.0 |
| IF-PC2 | 0.4640 | 219.19% | 10 | — |
| I - PC3 | 0.8301 | 132.28% | 13 | 2.5 |
| IF-PC3 | 0.4569 | 222.96% | 13 | — |
| I - RR1 | 0.8293 | 105.24% | 11 | 3.3 |
| IF-RR1 | 0.2434 | 220.36% | 8 | — |
| I - RR2 | 0.8325 | 101.45% | 10 | 4.4 |
| IF- RR2 | 0.2709 | 218.13% | 10 | — |
| I - RR3 | 0.8301 | 111.35% | 13 | 3.2 |
| IF-RR3 | 0.2363 | 222.31% | 9 | — |
| I - RR2A | 0.7796 | 16.50% | 8 | 4.4 |
| IF-RR2A | 0.2404 | 22.06% | 10 | — |

The poor showing of the IF set of models, generated by using the combined data from the States of Indiana and Florida, underscores the large differences in highway effects on land value for the two States (see Figure 24 ).

On the other hand, multivariate statistical analyses of models in the I set strongly substantiated the existence of a functional relationship between the mean impact of highway improvements on land value and the explanatory variables used in this research.

## Evaluation of Models in the I Set

Models in the I set that used DDIFF (difference in unit land value in constant dollars before and after highway improvement) as their dependent variable yielded the highest values of $R^2$. In fact of all the models developed in this set, models I-CC3, I-CC2 and I-CC1 occupy the first three positions in that order of preference based on $R^2$ values.

The improvements of models I-CC2 and I-CC3 over model I-CC1 are of special significance. At the preliminary stage of this work, scatter diagrams were drawn to get some insight into the relationships between the dependent variable (DDIFF) and two of the independent variables (size and time). These scatter diagrams revealed that these relationships were non-linear. However to include these variables into a multiple linear regression model these relationships must be linear. To better achieve linearity, model I-CC2 used logarithm transformations of these two independent variables. These transformations yielded some improvements over results obtained for model I-CC1. Model I-CC2 had an increased $R^2$ value over that of model I-CC1 by about one percent

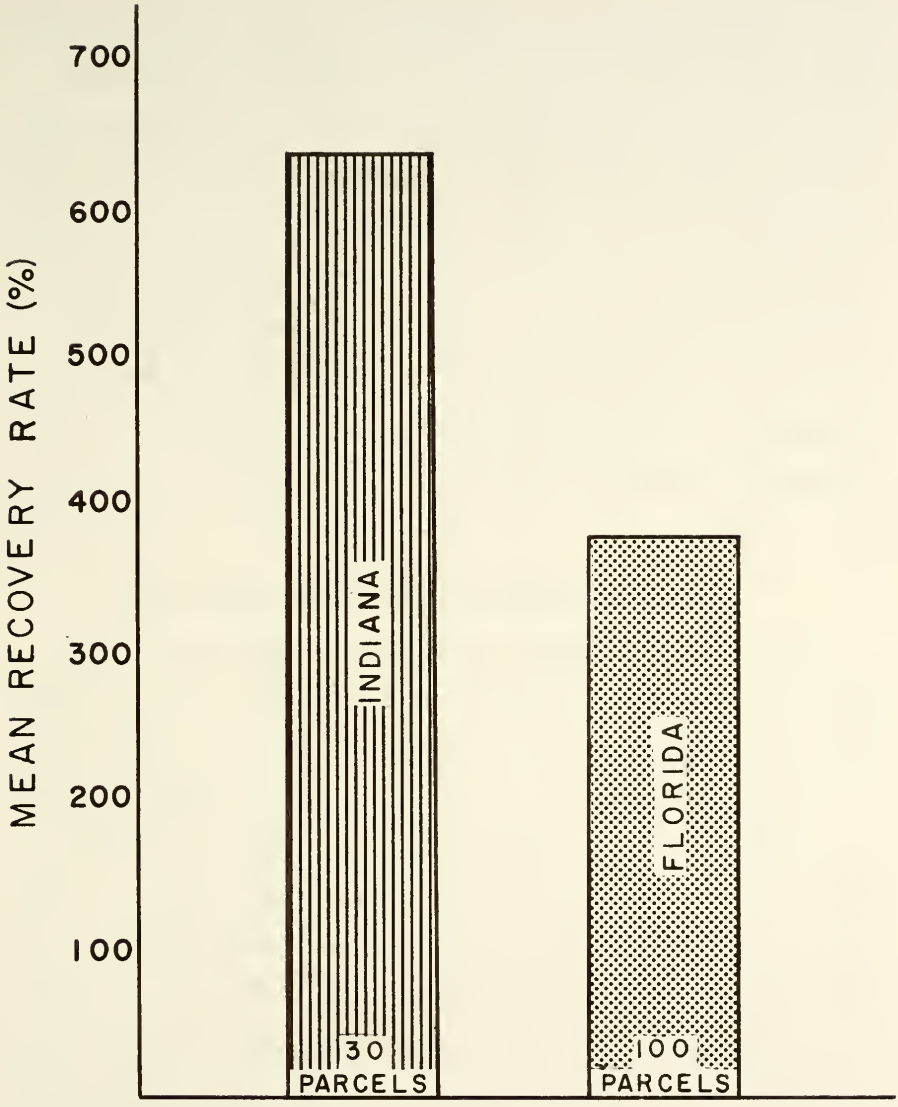



FIGURE 24: COMPARISON OF MEAN RECOVERY RATES FOR PARCELS IN FLORIDA AND INDIANA

and had a lower coefficient of variation of about four percent.

However when curvilinearity exists in any given relationship the task of obtaining the most appropriate transformation is not an easy one. The researcher may indeed have to try several transformations before deciding on one. The use of the dummy variable technique as discussed in chapter four provided a more direct and perhaps a more preferable way of handling this problem of curvilinearity. The use of dummy classes to represent the independent quantitative variables, size and time, in model I-CC3 yielded some encouraging results. Model I-CC3 produced the highest value of the coefficient of multiple determination ($R^2 = 0.8725$).

Models Using PDIFF or RVALUE as Dependent Variable

Models using PDIFF as the dependent variable produced, for the same set of input, results that were similar to those yielded by models using RVALUE as the dependent variable. This similarity of results is evident from Table 25 . This similarity is however not surprising; it might have been expected from the definitions of the dependent variables in Chapter Five. However, models using RVALUE as the dependent variable provided lower values of the coefficient of variation. Therefore, on the basis of the coefficients of variation and multiple determination, RVALUE was statistically preferable to PDIFF as a measure of highway economic impact on land value. Also unlike PDIFF, RVALUE can only assume positive values; thus it is amenable to a logarithm transformation.

## Changes In $R^2$ And In Standard Error of Estimate

An examination of Tables 5 to 14 reveals that most of the variation in the dependent variable is usually explained by the first seven or eight independent variables. Similarly, the standard error of estimate continued to decrease with the addition of more variables until about seven or eight variables were included; then addition of more variables usually led to an increase. Determination of where to stop adding more variables is a decision problem in which any additional statistical strength gained must be evaluated against the additional cost for data procurement and analysis.

## Predictive Strengths of Models in the I Set

The fact that a regression equation is statistically significant does not guarantee that it is a good predicting equation. A statistically significant regression equation at a specified risk level $\alpha$ only implies that the proportion of the variation observed in the data which is accounted for by the estimating equation exceeds what would be expected by chance in 100 $(1-\alpha)\%$ of the time. "Unless the range or values predicted by the fitted equation is considerably greater than the size of the random error, prediction will often be of no value even though a significant F-value has been obtained since the equation will be fitted to the errors only" ( 1 ).

A criterion for evaluating the adequacy of a regression equation for predictive purposes resulted from a dissertation by Wetz ( 2 ). According to his findings, an estimating regression equation can serve as a satisfactory predictor if the observed F-ratio is about four times

the selected percentage point of the F-distribution. This criterion was applied to models in the I set to test their usefulness for predictive work.

An $\alpha$ value of 0.05 is usually used in engineering work as the risk level. Therefore in reading the critical F-values ($F_c$) from statistical tables, a risk level of 0.05 was used.

The ratios of F values ($F_o$) observed from each I-set model to $F_c$ were computed and are tabulated in Table 25 . This model shows that six of the ten I set models satisfy Wetz's "four times" rule. On this basis, these models merit further considerations as candidates for the "best" model.

Further indications about the usefulness of these models are provided by Figures 25-30. These Figures represent plots of observed against estimated values of the dependent variables. They provide a quick view of how well the models predicted and are useful in identifying analysis areas which have unique characteristics. Elimination of such unique areas from consideration might help to improve the strength of developed regression equations. The models depicted in Figures 25-30 did not reveal the existence of any unique analysis area. This observation resulted from the even distribution of points around the forty-five degree line.

## Selecting the "Best" Regression Equation

In selecting the "best" model from a set of models, it is essential that the analyst considers both statistical and non-statistical criteria. However since the same guidelines governed the building of all the models in the I set, selection of the final model was based primarily on the

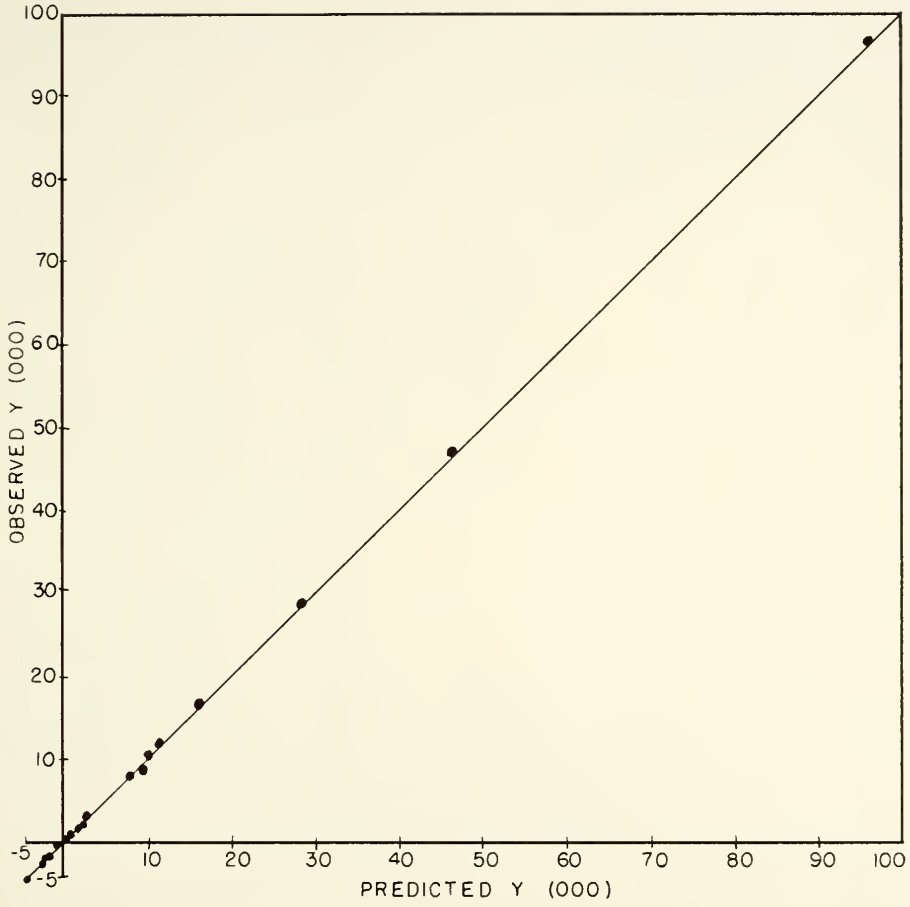



FIGURE 25: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL I-CCI

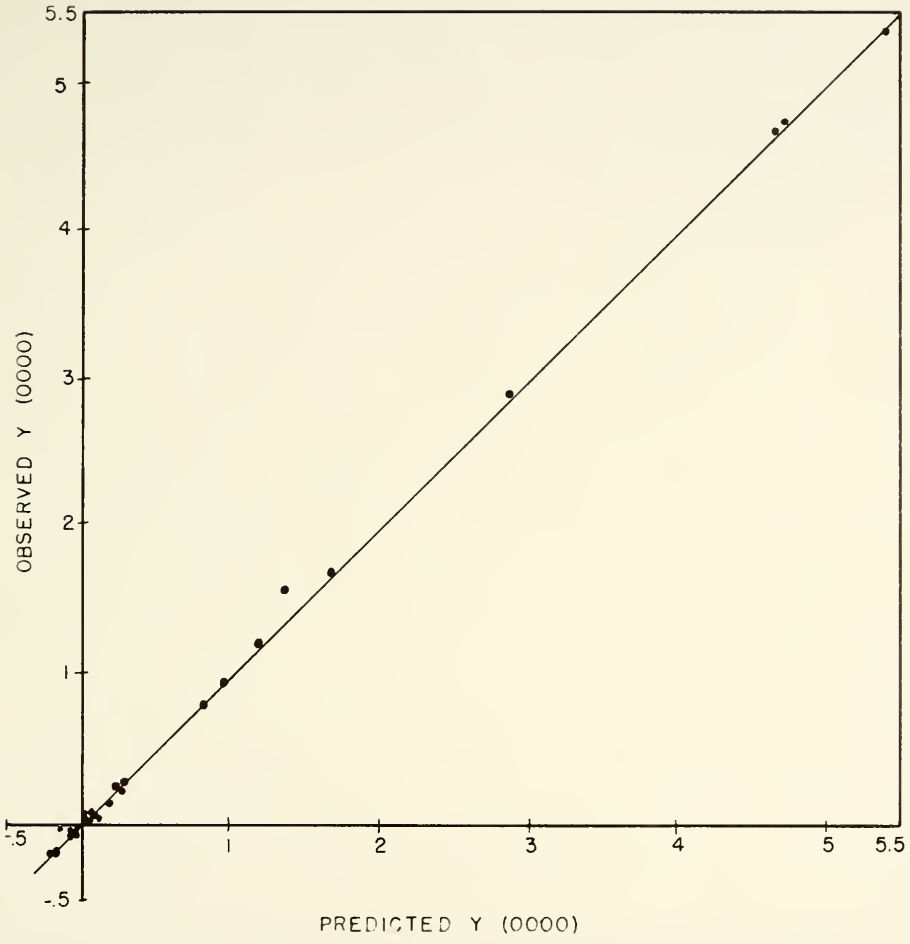



FIGURE 26: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL I-CC2

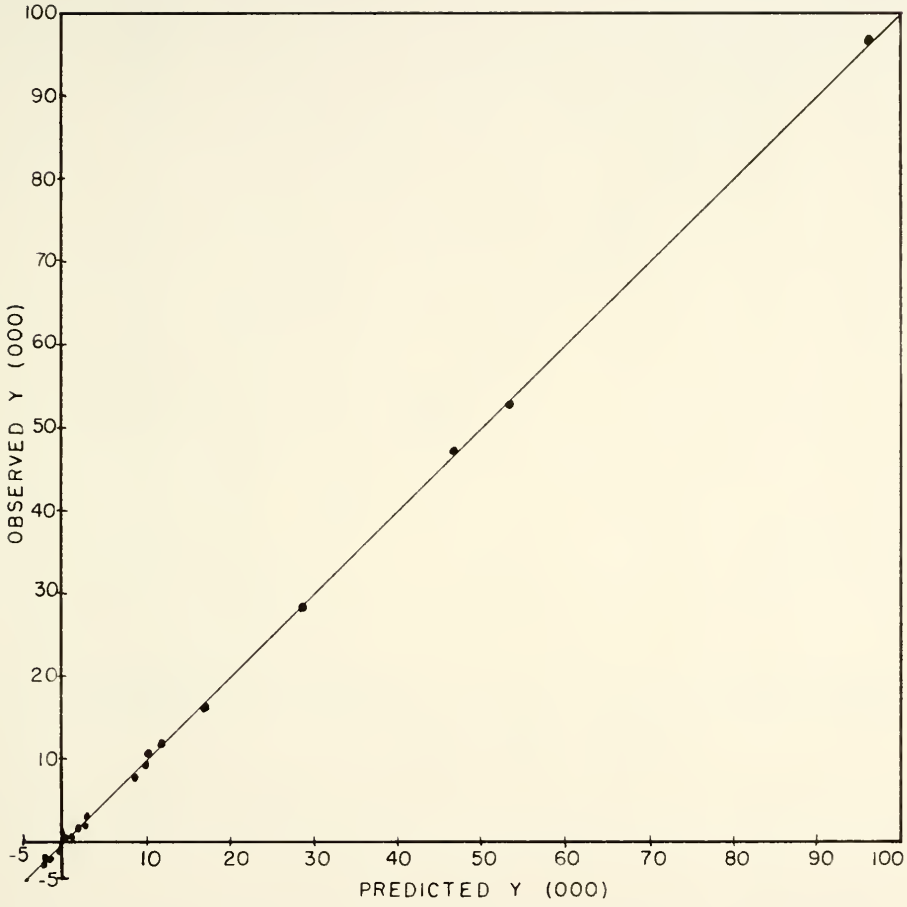



FIGURE 27: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL I-CC3

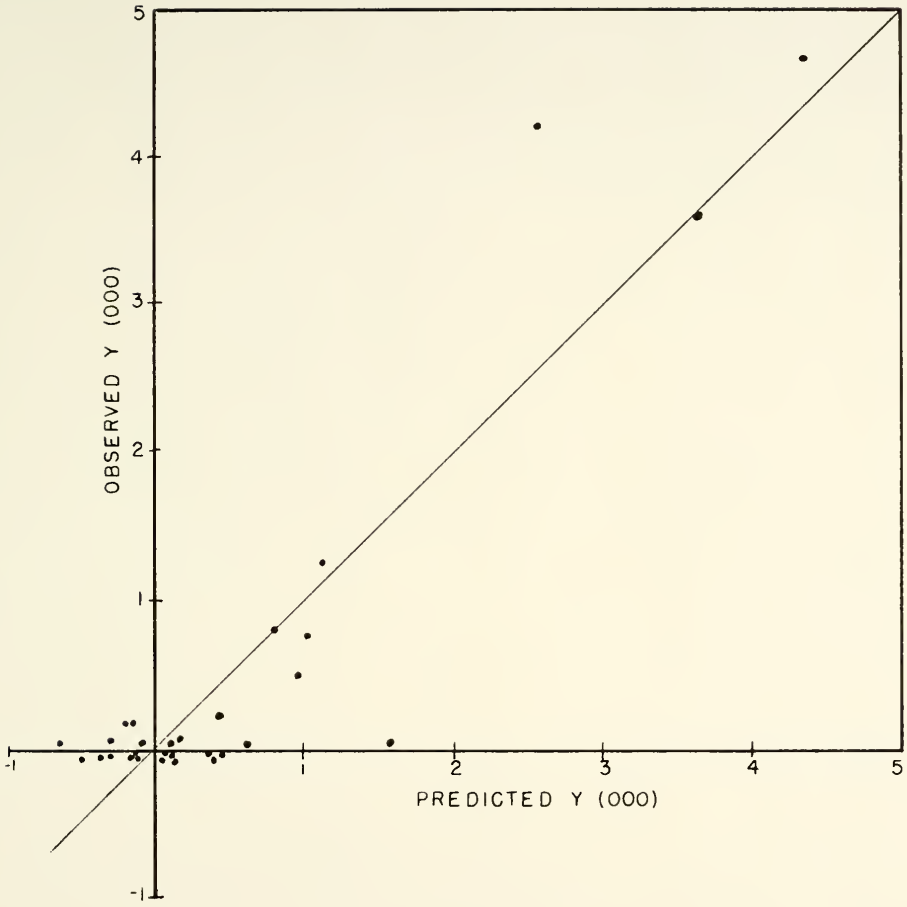



FIGURE 28: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL 1-PC2

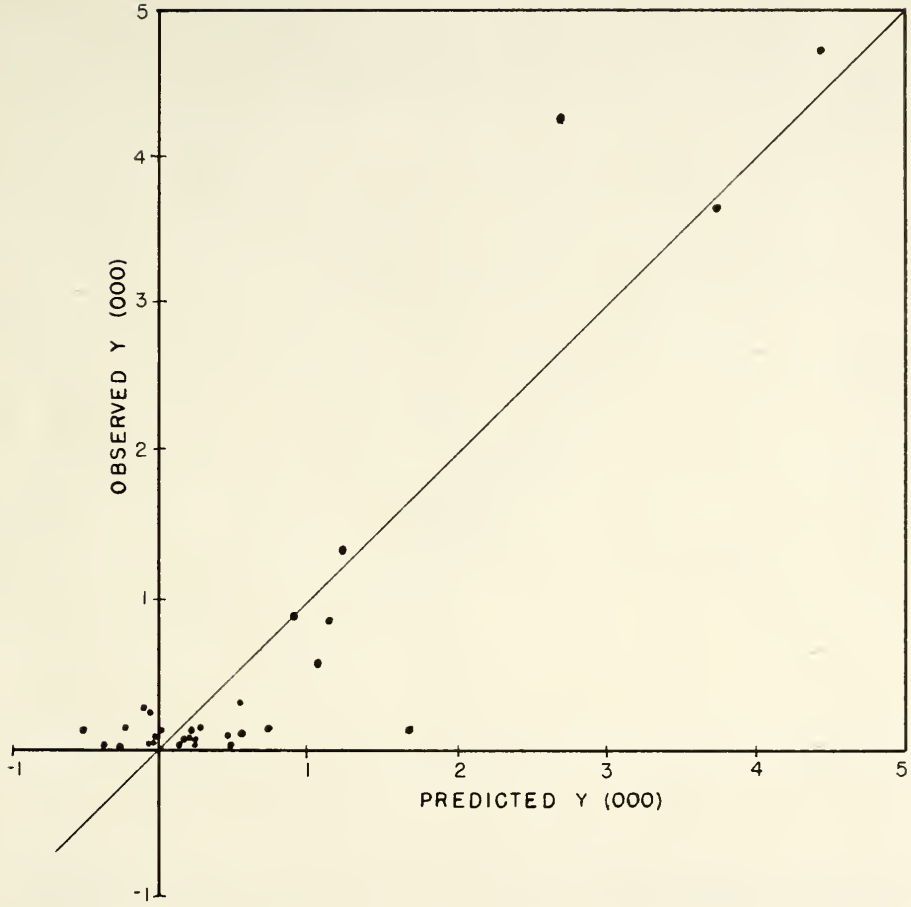



FIGURE 29: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL 1-RR2

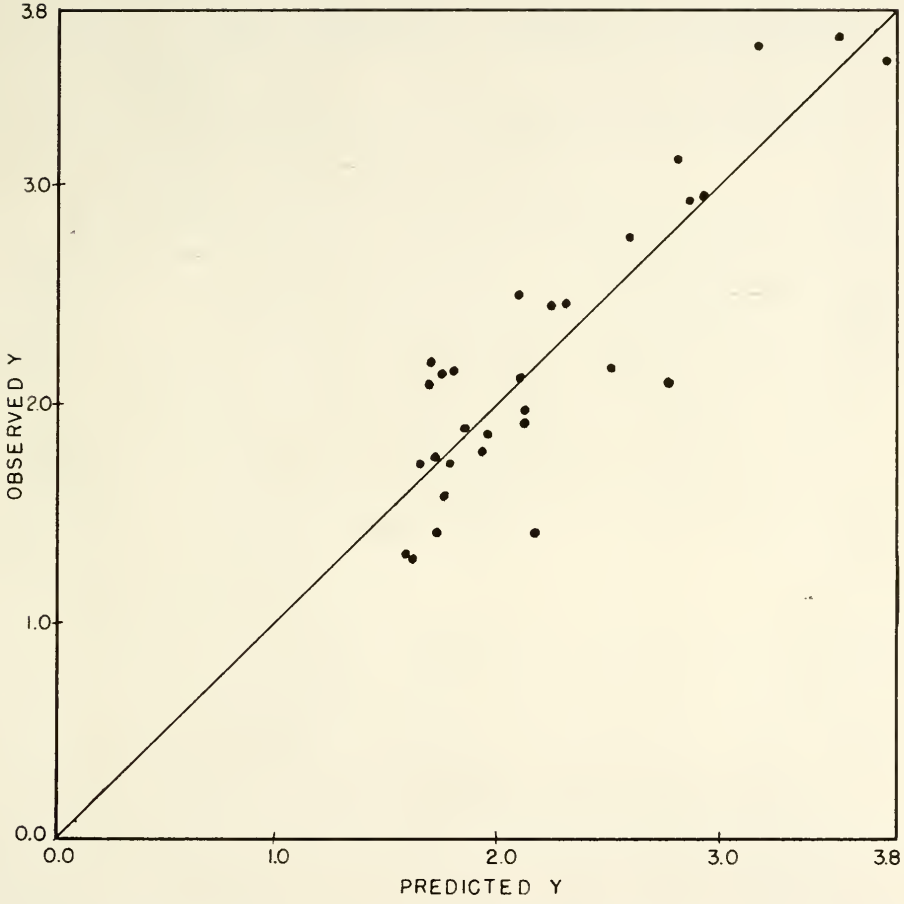



FIGURE 30: PLOT OF OBSERVED AGAINST PREDICTED VALUES OF THE RESPONSE VARIABLE FOR MODEL 1-RR2A

statistical performance of each model. Ideally an impact model is desired that has the following qualities:

1) contains those variables which are logical in terms of highway impacts
2) has as high a coefficient of multiple determination ($R^2$) as possible
3) has as low a coefficient of variation (C.V.) as possible
4) contains as few variables as possible

Values of $R^2$, C.V. and the number of independent variables included in each model are summarized in Table 25 . This Table shows that among the I set of models, model I-CC3 has the highest $R^2$ value and model I-RR2A the lowest. It is also worth-noting that both the second and third best in the order of $R^2$ values used the same dependent variable as model I-CC3. On the other hand, model I-RR2A has the lowest C.V. value (C.V. = 16.5%) while model I-CC3 has a relatively larger C.V. of 101.95% at its ninth step.

The problem of selecting the best model therefore reduces to making a decision on either of models I-RR2A or I-CC3. However the only consideration model I-RR2A has in its favor is its low coefficient of variation. Model I-CC3 not only has the highest coefficient of multiple determination but it also is easier to work with than model I-RR2A. This is so because model I-CC3 uses DDIFF (change in unit land value expressed in constant dollars) as its dependent variable. One does not encounter the computation involved in preparing the data for the dependent variable logarithm of the Recovery Rate) used by model I-RR2A.

The fact that model I-RR2A uses a fewer number of variables does not

necessarily make it preferable to model I-CC3. Table 7 shows that if model I-CC3 were terminated at the eighth step its $R^2$ value would be 0.8534 compared to 0.7796 for model I-RR2A, which used eight variables. In fact with only five variables in model I-CC3 its $R^2$ value was 0.7781.

Therefore on the basis of its high coefficient of multiple determination and the ease with which its input data can be prepared, model I-CC3 is probably the best overall model for the Indiana data.

## Relative Importance of the Independent Variables

The use of beta coefficients provided a means for measuring the relative importance of the individual independent variables in determining the dependent variable. Values for beta coefficients computed for the independent variables used in this research are tabulated in Table 26 for each of the developed models of the I-set. These coefficients indicate the expected increases in the dependent variables resulting from an increase of one standard deviation in each independent variable.

An examination of Table 26 shows that for each model the variable denoting highway type had the highest beta coefficient; indicating that this variable was the most important of all the independent variables considered. Following this variable (highway type) in order of importance for the Indiana data, were type of area, land use type, type of access control, time after the construction of the highway improvement and size of parcels.

Because of the limited data available for the development of the models in the I-Set, it cannot be stated that the relative importance of the independent variables to changes in land value is in general as quantified in Table 26 . The findings of Table 26 , however, are

# TABLE 26: VALUES OF BETA COEFFICIENTS FOR VARIABLES IN EACH MODEL

VARIABLES

| MODEL IDENTIFICATION β | HIGHWAY TYPE | TYPE OF AREA | TYPE OF LAND USE | TYPE OF ACCESS | TIME | SIZE |
|---|---|---|---|---|---|---|
| I-CC1 | 0.56 | 0.55 | 0.39 | 0.34 | 0.10 | 0.05 |
| I-CC2 | 0.61 | 0.56 | 0.37 | 0.33 | 0.13 | 0.12 |
| I-CC3 | 0.58 | 0.57 | 0.35 | 0.32 | 0.17 | 0.06 |
| I-PC1 | 0.66 | 0.60 | 0.57 | 0.27 | 0.12 | 0.04 |
| I-PC2 | 0.70 | 0.58 | 0.54 | — | 0.13 | 0.03 |
| I-PC3 | 0.65 | 0.60 | 0.63 | 0.32 | 0.12 | 0.03 |
| I-RR1 | 0.66 | 0.60 | 0.55 | 0.32 | 0.13 | 0.04 |
| I-RR2 | 0.70 | 0.58 | 0.49 | — | 0.13 | 0.03 |
| I-RR3 | 0.71 | 0.60 | 0.63 | 0.30 | 0.11 | 0.02 |
| I-RR2A | 0.71 | 0.57 | 0.48 | 0.32 | — | 0.27 |

indications of the relative importance to changes in land value of the independent variables. Highway type, type of area and type of land use appear to be most important with type of access and time of sale of lesser importance and size of tract relatively unimportant.

### Relative Importance of Classes of Highway Type

Figure 22 shows that from the Indiana data land parcels along secondary routes received more enhancement than parcels either along primary or interstate routes. The latter group of parcels, in fact, experienced very little increase in land value. However this result should not be interpreted to mean that secondary and primary routes always experience more enhancement than those on interstate routes. A probable reason for these results in this analysis might be because most of the existing interstate mileage was in rural areas at the time the data used in this work were obtained (Reference 3 p. 78). Moreover only nine of the parcels considered here were situated either along secondary or primary routes and they might not be truly representative of their classes.

### Relative Importance of Types of Area

Land parcels located in urban areas experienced much greater benefits from highway improvements than did parcels located in either the fringe or rural areas. This is also shown in Figure 22 for Indiana data. The number of urban parcels included in the data, however, was small and the very large increase in value of land in this type of area may not be truly representative. It is probable, however, that increases in

land value are small in a rural area and much greater in and near urban areas. Figure 31 also shows this to be true for the large amount of data from Florida. Here increases in land value of urban property adjacent to the highway improvement was about six times that of rural property.

## Relative Importance of Classes of Land Use

Figure 22 shows that for the Indiana data parcels whose land uses were denoted either as commercial or vacant at the time of acquisition were more enhanced in value than properties that were used for residential and agricultural purposes. However, only one and two parcels respectively were available for commercial and vacant land uses, and the land value changes found may not provide a true picture of the mean effects.

The increase in land value was also found to be much greater for residential land use than for agricultural land use, possibly an indication of a greater density of development, where one would expect demand for land to be greater. In both cases much of the increase in land value is probably due to an anticipated change in land use, such as from agricultural or residential to commercial.

Data on land use type from Florida indicates that commercial land suffered a loss in land value while residential, vacant and agricultural increased (see Figure 31 ). The increase in value of land used for residential purposes was much greater than that for agricultural and vacant land.

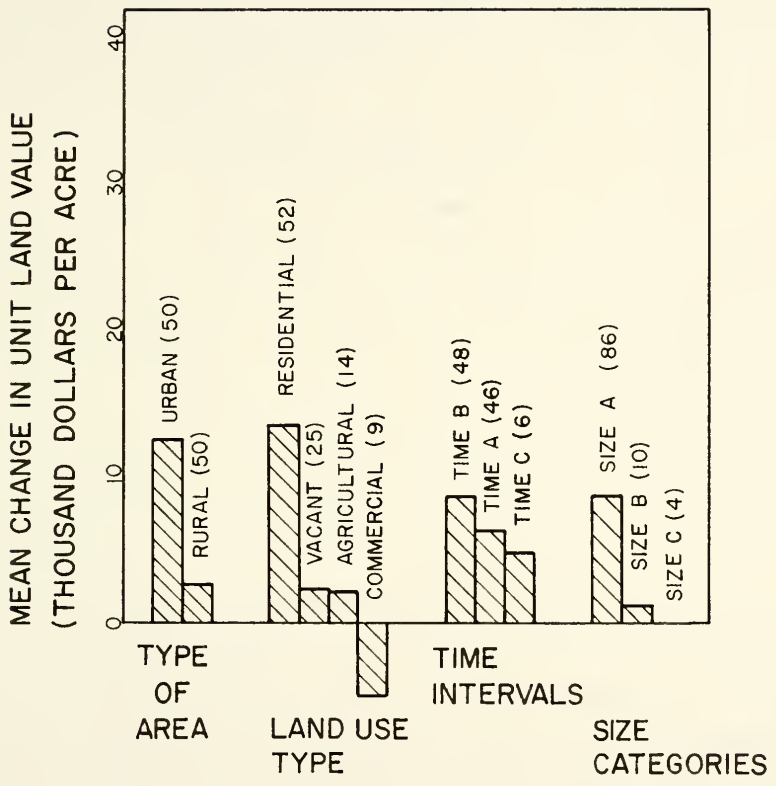



FIGURE 31 : MEAN CHANGE IN LAND VALUES FOR CLASSES WITHIN EACH VARIABLE

## Relative Importance of Types of Access

The type of access existing between a highway and parcels along it appears to affect the recorded changes in land value. Figure 23 from the Indiana data reveals that parcels with no access or only partial control of access to the highway benefited most. In fact parcels whose direct access to the highway was completely denied suffered a small loss in mean unit value.

## Importance of Time to Land Value Changes

Figure 23 of Indiana data reveals that increases in the mean value of land resulting from the highway improvement are experienced during an initial period of 18 months but are greater for longer periods. Maximum increase appears to occur during the period of 18 to 41 months after completion of the highway improvement. The increase in value, however, for longer periods of time was also greater than that for the initial period.

The data from the 100 parcels along the interstate system in Florida also indicate maximum increase in land value occurs during the 18-41 month period (see Figure 31 ). These data, however, show that increases in land value after 42 months are slightly less than during an initial 18-month period. It should be noted that this latter indication is based on only six land parcels.

## Relative Importance of Size Categories

Small parcels received greater increases in land value than did large parcels. In fact some large parcels experienced some loss in value. This is revealed by the Indiana data in Figure 23 where parcels in category SA (less than fifteen acres in size) experienced the greatest increase in value.

The Florida data, shown in Figure 31 , provides similar information. Small parcels increased much more than medium size parcels and large parcels (over 50 acrea) suffered a slight loss in mean value.

# CHAPTER EIGHT

## DEVELOPMENT OF A NON-DETERMINISTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM: A CONCEPTUAL FRAMEWORK

### The Impact Problem Revisited

The regression models developed in this research provide a deterministic approach to solving the problem of predicting land value impacts resulting from highway improvements. They are deterministic in the sense that given a set of values of the independent variables, each of these models gives one and always the same value for the response variable. They therefore provide a one-shot estimate of what will happen to the average land parcel under a given situation. Thus the problem of predicting the economic impacts of highway improvements on land value may be made to appear much more simplified than is actually the case.

Experience and knowledge about the impact problem situation suggests, however, that the problem is non-deterministic. The impact problem is actually surrounded by many uncertainties that are so significant to the problem solution that they cannot be ignored. Changes in population, customer purchasing habits, personal disposable incomes, car ownership and the fluctuations of the stock market represent only a few of the dynamic forces that influence land values. Under such circumstances it is very difficult, if not impossible, to single out with certainty a direct causal relationship between a given highway improvement and an

observed change. The Bureau of Public Roads summed up this uncertainty very nicely in these words:

> "....there is no definite certainty that a particular type of highway improvement will lead inevitably to a particular set of consequences. For example, no assurance can be given that widening a 2-lane highway to a 4-lane facility will cause traffic flows of a specific number of vehicles or land value increases of a definite amount." ( 1 )

Regression models are designed to predict the mean effects. They provide us with estimates of what will likely happen to the average land parcel. There might however be wide variances of many land parcels from the imagined "average" parcel. One might desire a much deeper insight into the actual impacts occurring after highway improvements than the limited indication provided by observing the average impact. In addition, several assumptions are made in using the regression technique. In tests for the significance of the regression model, the normality assumption is made. In multiple linear regression the assumption is made that the relationship being described is one where the variables are purely additive. These assumptions are not always true and may lead to an error in our estimates of the dependent variable.

To meet these weaknesses of regression models, some ground work was undertaken, and is reported in this chapter, to provide a non-deterministic approach to the land value impact problem. This approach is based on the premise that uncertainty is inevitable in developing a predictive technique for land value impacts. One can, however, reveal a fuller spectrum of impact states that may occur following a given

highway improvement and provide some probabilistic descriptions about the occurrence of each state. Working within the principles of probability theory and examining empirical data, such probabilistic statements can be evolved.

In developing this methodology, an apparent analogy between the impact problem and the medical diagnosis problem was first examined.

## An Analogy Between The Impact Problem And Medical Diagnosis

Certain resemblences seem to exist between the problem of predicting the economic impacts of highway improvements on land value and the medical problem of diagnosing a patient's disease on the basis of some observable symptoms. A skillful land appraiser or an engineer experienced in real estate fluctuations have some indicators about a land parcel which enable him to predict the probable impact of a given highway improvement on its land value. Although he might not be able to put numerical values on the likelihood of a given land parcel going through an impact state $R_j$ he can subjectively ascribe his degree of belief as to this in terms of the frequency with which he has observed this in the past. The knowledge from which he retrieves this subjective evaluation of the impact has been acquired through years of actual experience. These indicators can be likened to the symptoms which doctors look for in medical diagnosis.

A close look at medical diagnosis reveals that it is a process of pattern matching. A doctor draws on available information (patient's history and observed symptoms) about a patient and them forms a mental picture of the likely disease pattern, and with a recall from his memory

tries to match this pattern with known patterns. To be fast and efficient at diagnosis, a doctor must therefore have a rich storage of symptom-disease complexes. However the ability of the human brain to be an efficient storage and retrieval system becomes doubtful when the number of these symptom-disease patterns becomes large. The development of computers has greatly increased the number of possibilities that can be tested and has provided a more reliable means of matching similar patterns.

Hoffer (2) identified two approaches in which computers can be applied to medical diagnosis as:

1. The logic tree approach
2. The probabilistic approach

The logic tree approach attempts to trace the path followed by a doctor during his decision-making process after he has observed some symptoms or has examined the results of pathological tests. This approach gets its name from the various decision points involved in the process; these decision points can be represented graphically as a series of branching paths from a tree. In the probabilistic approach, statistical formula such as the Bayes' theorem are applied to provide various diagnostic possibilities; each with the probability of its occurrence.

An article (3) by Ledley and Lusted in 1959 is credited as the first on the subject concerning the use of Bayes' theorem of conditional probability for computer-aided diagnosis. Warner (4) however was the first to apply this statistical theorem on real data.

Suppose we have a set of symptoms $\vec{S}$ and a set of diseases $\vec{D}$ where

$$\vec{S} = (S_1, S_2, S_3, S_4, \ldots, S_m)$$

$$\vec{D} = (D_1, D_2, D_3, D_4, \ldots, D_k)$$

Then a doctor confronted with a patient who has certain symptoms $\vec{S}_j$, would like to assert with some probability that the patient has a disease $D_i$. In other words, we want to know p $(D_i/\vec{S}_j)$ which is the mathematical way of stating the problem. However what is prescribed in all medical textbooks is p $(\vec{S}_j/D_i)$, that is the probability of having the symptom complex $\vec{S}_j$ when it is known that the patient has a disease $D_i$. Although these textbooks presently do not give numerical values of P $(\vec{S}_j/D_i)$, they do stipulate that associated with a certain disease, particular symptoms are known to be common, frequent, rare, etc. Reference is usually made to P $(\vec{S}_j/D_i)$ rather than P $(D_i/\vec{S}_j)$ for the simple reason that the etiology of the symptoms is associated with or caused by the disease. Thomas Bayes, an English Philosopher, provided a formula which enables a physician, for example, to transfer what he learned in medical schools as to the values of P $(\vec{S}_j/D_i)$ to the actual problem he faces when a patient's disease is to be diagnosed. In general terms, Bayes' formula states that if we have N mutually exclusive and exhaustive events $C_1, C_2, \ldots, C_n$ and an event B for which we know the conditional probabilities $P(B/C_i)$ of B, given that $C_i$ has occurred and also the absolute probabilities $P(C_i)$, then the conditional probability of $P(C_i/B)$ of any one of the events $C_i$ given B is given by the formula:

$$P(C_i/B) = \frac{P(B/C_i)\ P(C_i)}{\sum_{j=1}^{n} P(B/C_j)\ P(C_j)} \qquad \text{---- (8-1)}$$

Therefore if we call the events $C_i$ "Causes", Bayes' formula can be interpreted as providing the probability that the event B, which has occurred is the result of the "cause" $C_i$ (5). The interested reader is referred to reference (6) for the details of Bayes' original essay.

Warner, Toronto and Veasy have successfully applied Bayes' theorem to the diagnosis of thirty-three congenital hart diseases (7). Overhall and Williams (8) and Fitzegerald and Williams (9) also used this technique in the diagnosis of thyroid disorders. They found that a computer program based on Bayes' theorem successfully predicted 96% of the cases. Judging from these results, one can envisage greater usage of this approach in medical diagnosis.

One of the operational mechanisms used in evolving a solution in a problem-solving situation is to make a direct analogy between the problem at hand and a similar phenomena in other disciplines. In this regard, a one to one correspondence may be hypothesized between the medical diagnosis problem and the problem of predicting the effects of highway impacts on land value. Based on this premise a conceptual framework will now be described for coping with the impact problem.

## Outline Of Bayesian Approach To The Impact Problem

In this development of the Bayesian approach to the impact problem, recovery rate (RR) as defined in Chapter Five will be used as the measure of the response of land value changes to highway improvements.

Let $F_1$ (a), $F_2$ (b), ....., $F_N$(k) represent a set of independent and fundamental factors explaining most of the variations in the observed response. The subproblem of deriving these basic factors will not be considered at this stage but deferred to a latter stage. This set of factors can then be used as the independent variables in a discriminant function which classifies each land parcel into a category depending on some boundary values. Then a land parcel in category j denoted by $C_j$ can be described as:

$$C_j = (F_1(a_j), F_2(b_j), F_3(d_j), \ldots F_N(k_j))$$

If a comprehensive analysis of historical data of parcels that are known to have belonged to category $C_j$ reveals that parcels in this category experienced different states of highway impact that could be defined by:

$$\vec{R} = (R_1, R_2, R_3, \ldots, R_m)$$

then each element $R_j$ represents a state of highway impact. Each such impact state may be definable in terms of recovery rates. It can be assumed that the relationship between recovery rate and time has a constant slope from time t=0 to some other time t=$t_p$ and that for any time greater than $t_p$ the slope is zero. The land parcels in category $C_j$ then can experience impact state $R_1$, where $R_1$ represents the range $(R_L - R_U)$

where $$R_L = \frac{(RR_L)_{to} + (RR_U)_{to}}{2}$$

and $$R_U = \frac{(RR_L)_{tp} + (RR_U)_{tp}}{2}$$

$(RR_L)_{tp}$ and $(RR_U)_{tp}$ are the lower and upper values of the recovery rate of the parcels in this state at time t=tp. Also $(RR_L)_{to}$ and $(RR_U)_{to}$ represent these values at time t=0.

The frequency with which parcels in category $C_j$ fall into any state can be computed from past records. As in other predictive efforts, one must often assume that the trends of the past also provide a reflection of possible future occurrences. Then probabilities computed on the basis of past and present events can be used to represent the probable impact of highway improvements on land value in the future.

The problem of predicting the impacts of highway improvement on the value of a land parcel can then be stated simply as follows: Given, the description of a land parcel in terms of some basic factors, what is the probability of the parcel being in an impact state $R_j$? Discriminant functions developed from past known impacts can evaluate the descriptors of this land parcel and on the basis of this evaluation assign it to a category $C_k$. Then the impact problem reduces to computing the probability $P(R_j/C_k)$.

The problem is thus reduced to a form in which Bayes' formula can generate the solution. This formula is

$$P(R_j/C_k) = \frac{P(R_j)\ P(C_k/R_j)}{P(C_k)}$$

Now consider the meaning of both the denominator and the numerator of the right hand side of this formula. Suppose there are n states of impact for parcels in category $C_k$. Then from the above equation:

$$P(R_j/C_k) = \frac{P(R_j)\ P(C_k/R_j)}{\sum_{\theta=1}^{m} P(R_\theta)\ P(C_k/R_\theta)} \qquad (i=1, 2, \ldots n)$$

For i = 1

$$P(R_1/C_k) = \frac{P(R_1)\ P(C_k/R_1)}{\sum_{\theta=1}^{m} P(R_\theta)\ P(C_k/R_\theta)}$$

For i = 2

$$P(R_2/C_k) = \frac{P(R_2)\ P(C_k/R_2)}{\sum_{\theta=1}^{m} P(R_\theta)\ P(C_k/R_\theta)}$$

---

For i = m

$$P(R_m/C_k) = \frac{P(R_m)\ P(C_k/R_m)}{\sum_{\theta=1}^{m} P(R_\theta)\ P(C_k/R_\theta)}$$

$$\sum_{j=1}^{m} P(R_j/C_k) = \frac{\sum_{j=1}^{m} P(R_j)\ P(C_k/R_j)}{\sum_{\theta=1}^{m} P(R_\theta)\ P(C_k/R_\theta)}$$

$$= 1$$

Thus the denominator of equation ( 8-1) is a normalizing factor serving to make $\sum_{j=1}^{m} P(R_j/C_k)$ equal to 1. This also implies that in comparing

the probabilities of parcels in category $C_k$ in two of the m states of impact, it is not necessary to compute the denominator and comparisons can be done on the basis of the magnitudes of the respective numerators.

The term $P(C_k/R_j)$ in the numerator is the probability of membership in category $C_k$ when it is known that the parcel is in impact state $R_j$ after the highway improvement.

The second term $P(R_j)$ in the numerator is the probability that the particular population of land parcels in category $C_k$ experience an impact state $R_j$. This term has to be defined for a specific sample space and therefore recognizes differences that might exist in the magnitudes of highway impacts from one geographical area to another. $P(R_j)$ is called the a priori probability of the occurrence of $R_j$ because it is known before observations are made.

On the other hand $P(R_j/C_k)$ is called the posterior probability as it is a conditional probability that is only of interest after the occurrence of an event.

The impact problem therefore reduces to observing the descriptors of a given land parcel and on the basis of past information determining the probability of its being in various impact states. The form of the reduced prediction problem is shown in Figure 32.

In the development of the above technique for predicting the effects of highway improvements through the use of Bayes' formula, two subproblems were encountered.

Subproblem 1: Given a large list of explanatory variables that are known or suspected to be associated with the impacts of highway improvements on land value, how can one evolve from this list a much

## THE PREDICTIVE PROBLEM

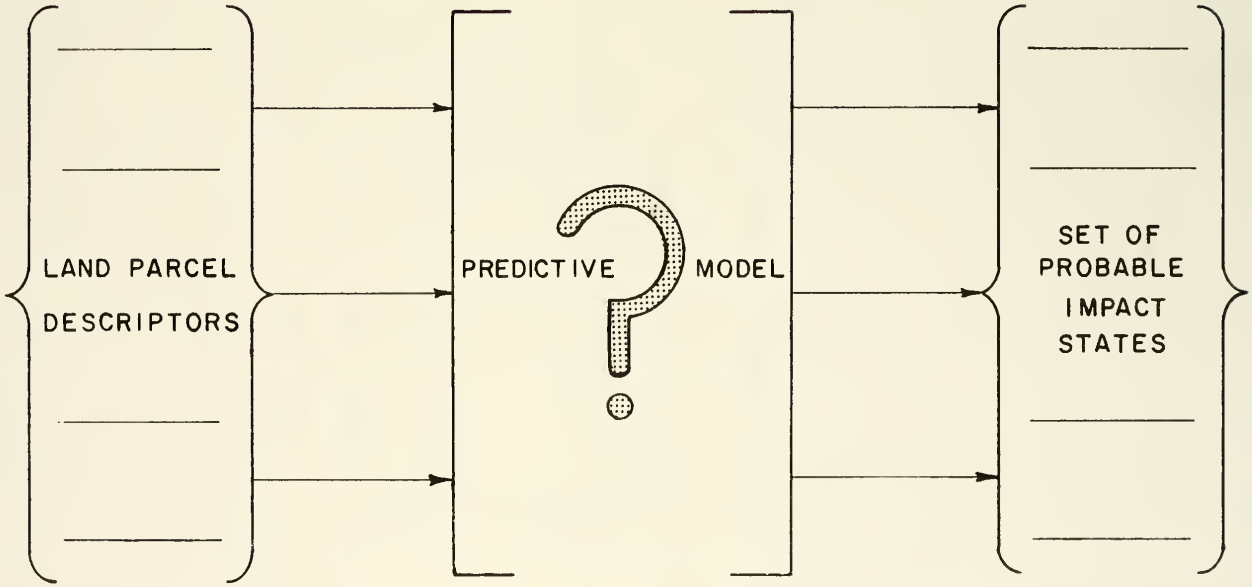



FIGURE 32: MAIN COMPONENTS OF A DIAGNOSTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM

smaller number of independent variables that best summarizes the original set of variables?

Approaches to problems of this kind in the past have centered around the use of the multiple regression technique. Dubois (10) proposed an iterative application of the multiple regression technique for eliminating some variables from a larger set. In the approach suggested by Dubois the variable possessing the highest beta coefficient at each stage is removed from further consideration. This cycle is repeated until the desired number of variables was obtained at the termination of this iterative procedure. The eliminated variables at the various stages then form the set of "basic" variables required to predict the response variable.

A modified version of the Dubois approach eliminates at each stage of the regression analysis the variable with the lowest beta coefficient. According to Wortman (11) this modified version was superior to the former in his studies because

1. the set of explanatory variables obtained by this modified version gave higher multiple correlation coefficients and lower standard errors of estimate when compared to an equal number of variables obtained by other methods.
2. this procedure compensated for instances in which a high correlation existed between some independent variables. Such a high correlation continued to reduce the beta coefficient until one of the interrelated variables was eliminated.

The correlation of the eliminated variable at each stage was checked in Wortman's work to insure that there was a logical basis for

the elimination. The criterion to be be satisfied for a logical elimination was stated as follows:

1. There was low correlation with the response variable,
2. There were high correlations with other variables, or
3. A combination of the above two.

Factor analysis offers a third and perhaps the best approach for tackling this subproblem. Factor analysis was first used in psychology and related areas to analyze the interrelationship between intelligence and personality tests (12). The principal objective of factor analysis is to attain a parsimonious description of observed data. A satisfactory application of factor analysis must therefore yield factors which convey all the essential information of the original set of variables. The researcher, using factor analysis to reduce the number of variables with which he has todeal might use one of two criteria:

1. To extract the maximum variance, or
2. To best reproduce the oberved correlations.

An empirical method for the reduction of a large body of data so that a maximum variance is extracted was first proposed by Karl Pearson and fully developed as the method of principal components by Harold Hotelling.

In contrast to the maximum variance approach, the classical factor analysis model is designed to maximally reproduce the correlations. The basic factor analysis model may be put in the form

$$Z_j = \sum_{i=1}^{m} a_{ji} F_i + d_j U_j \quad \forall j \ (j=1,2,\ldots, n)$$

in which each of the n observed variables is described in terms of m (usually smaller than n) common factors, $F_i$ and a unique factor $U_j$. The common factors account for the correlations among the variables while each unique factor accounts for the remaining variance (including error) of that variable. The coefficients of the factors are frequently referred to as factor loads.

A detailed description of the procedure used in factor analysis is given in reference 12. Moreover, several prepackaged computer programs of factor analysis are now available for use. One such program available on the CDC 6500 computer at Purdue University is the Miami Factor Analysis program. This program computes intercorrelations, principal component eigenvalues and factors loadings using the varimax method of rotation.

Subproblem 2: The second subproblem was a problem of classification. It can be stated this way: Given a set of variables characterizing some known land parcels that have been affected by highway improvements, how does one develop a set of rules for assigning any new land parcel to one of the classes of classification?

The method of discriminant analysis as developed originally by R. A. Fisher (13) is directed at this problem of classification. This technique employs a set of variables (in our case those resulting after factor analysis) characteristic of the problem at hand and sets out to find a linear combination of the variables that will permit an accurate allocation of objects into one of the classes of classification. The coefficients of the discriminant functions are then computed by using a procedure similar to that used in multiple regression analysis.

However while the least square criterion is used in regression analysis to compute regression coefficients, in discriminant analysis the "Mahalanobis distance" between the known groups, as defined by the difference in mean values for constant within group variance, is the quantity maximized in the computation of these coefficients.

Initially, discriminant analysis is used to classify a set of land parcels whose classes of classification are known. The number of errors of classification obtained in this process provides an insight into how well the developed discriminant functions would serve. These developed discriminant functions also have associated with them a set of boundary values which are used in the classification procedure.

An iterative discriminant analysis technique was developed by Casetti (14) in which discriminant procedures are repeatedly developed until an optimal classification and optimal procedures are obtained. Using his iterative technique, discriminant functions on land parcels that have been initially classified can be developed. These developed discriminant functions are then used to determine the probabilities of membership in various categories for each of the parcels used in their calibration. The probabilities of membership of some parcels in certain categories might be higher than those for the categories to which they belong. Such parcels are then reassigned to those classes in which they have the highest probabilities of membership and a new discriminant function is developed. This procedure is continued until each parcel has its highest probability in the category to which it was assigned prior to the calibration of the discriminant function. The classification of the parcels when this "end point" is reached is called the limit or

optimal classification. The discriminant function used in developing this optimal classification is then considered satisfactory to classify parcels other than those used in its calibration.

In situations where there are only two groups, R. A. Fisher suggested a test of significance which makes use of Hotellings generalized Student distribution. He defined a quantity $R^2$, analogous to the multiple coefficient of determination, as:

$$R^2 = \frac{N_1 N_2 (K_1 d_1 + K_p d_p)}{N}$$

Where

$N_1$ = number of objects in first group

$N_2$ = number of objects in second group

$N$ = $(N_1 + N_2)$

$d_i$ = difference of the means of variable i for the groups

To test the hypothesis that the empirical discriminant function may have risen out of pure chance if in reality there is no difference between the variates in the two groups, Fisher also defined a variance ratio as:

$$F = \frac{(N - p - 1) R^2}{p (1-R^2)}$$

which has Snedecors F distribution for $n_1=p$ and $n_2=N - p - 1$ degrees of freedom.

## A Sample Problem

To illustrate the methodology suggested in this chapter a simplified example problem will now be formulated and solved following the stages outlined by Figure 33.

### Problem Statement

A highway improvement is being contemplated for a given urban area. At the preliminary stage of planning the decision-makers are concerned about two land parcels, A and B described in Figure 34. They would like to predict what impact the proposed primary highway will have on the land values of these two parcels. The owners of these land parcels are influential men in the social and economic life of this community. As a result, the decision-makers are desirous to have a strong rational basis to substantiate predictions of future land values for parcels A and B; at least stronger grounds than mere off-hand predictions.

### Solution

#### Stage 1: Deriving Basic Factors

Assume here that these decision-makers had in the past familiarized themselves with the variables underlying impacts on land value for various highway decisions. To further simplify this sample problem, assume that their analysis of past data revealed that only two variables are basic to land value impact, namely: type of land use (LU) and distance from the highway improvement (DI).

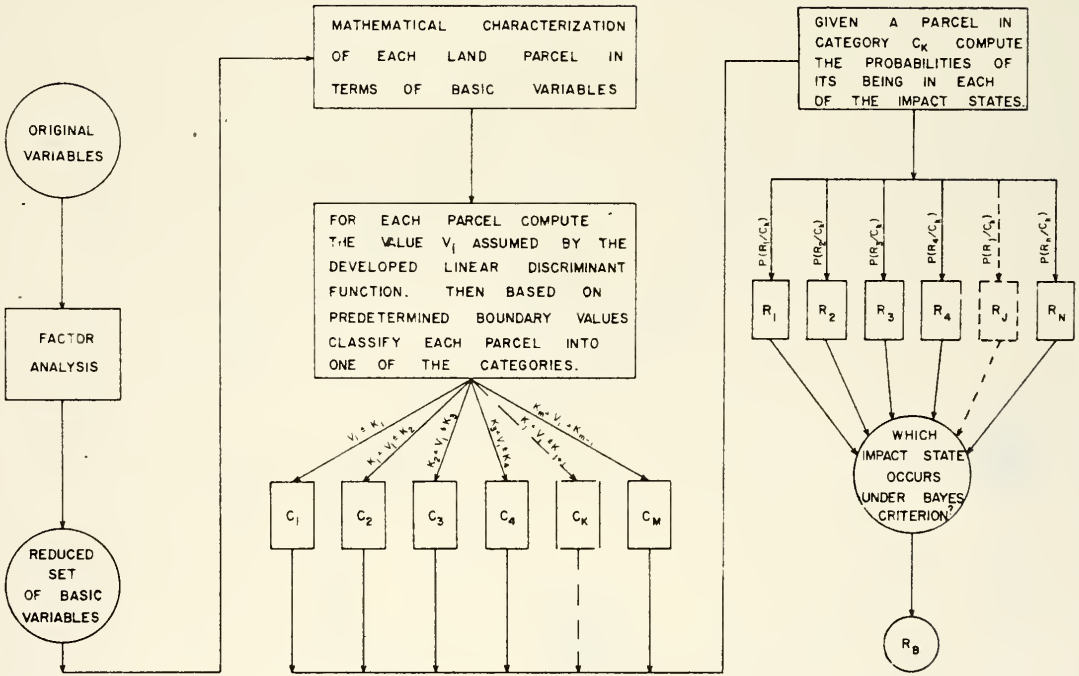



FIGURE 33: STAGES IN THE DEVELOPMENT OF A NON-DETERMINISTIC APPROACH TO THE IMPACT PREDICTIVE PROBLEM

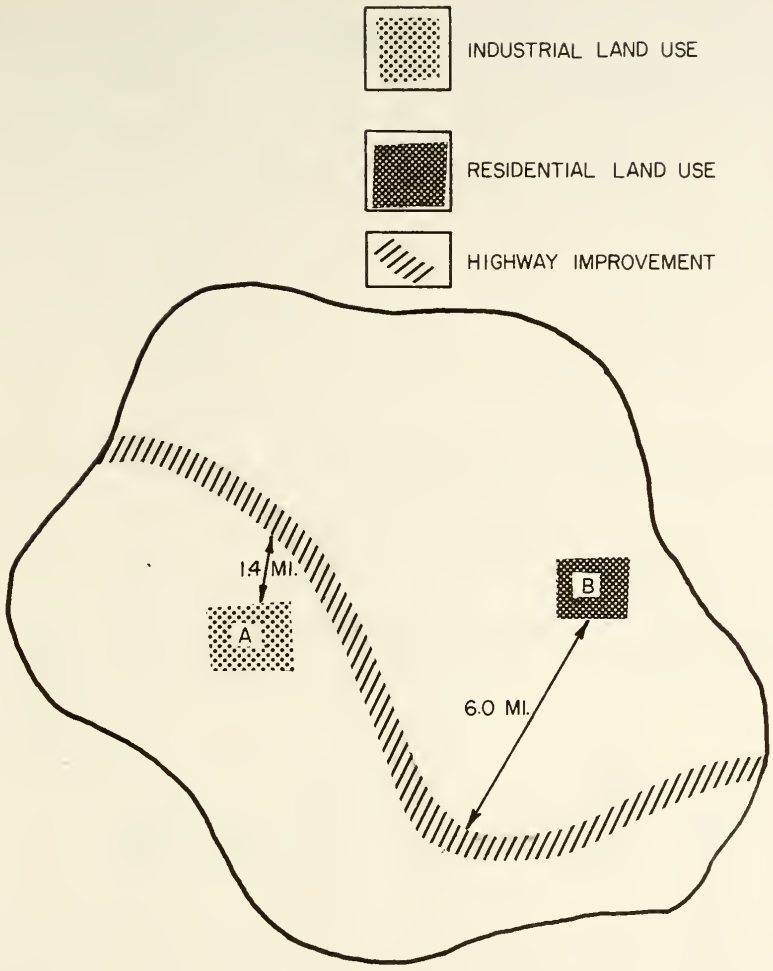



FIGURE 34: LOCATION OF PARCELS IN THE SAMPLE PROBLEM IN RELATION TO THE HIGHWAY IMPROVEMENT

Stage II: Classifying the Parcels into Categories

The following identification indices are used to describe various types of land uses:

| Land Use Type | Identification Indices |
|---|---|
| Industrial | 1 |
| Commercial | 2 |
| Agricultural | 3 |
| Vacant | 4 |
| Residential | 5 |

Then the land parcels A and B can mathematically be described as follows:

$$LP(A) \equiv \{LU(1),\ DI(1.4)\}$$

$$LP(B) \equiv \{LU(5),\ DI(6.0)\}$$

Further suppose that a discriminant function has been derived from an analysis of past data and that this function is given by:

$$F = 8 + 5X + 3Z_1 + 4Z_2 + 6Z_3 + 2Z_4 + 5Z_5$$

where x represents the value assumed by b in DI (b)
$Z_i$ (i = 1, ...5) are dummy variables each representing a type of land use and i is the identification index in LU(i). Each $Z_i$ takes a value of 1 or 0 depending on whether a given land use type is present or absent.

In conjunction with this function there are four boundary values which determine the membership of each land parcel in one of four categories. These boundary values and their associated categories are as follows:

| Range of F values | Category |
|---|---|
| $F < 10$ | $C_1$ |
| $10 < F \leq 20$ | $C_2$ |
| $20 < F \leq 35$ | $C_3$ |
| $F > 35$ | $C_4$ |

On this basis, for parcel A we have,

$$F_A = 8 + 5\ (1{\cdot}4) + 3\ (1) = 18$$

$\therefore$ parcel A is in category $C_2$

Similarly for parcel B

$$F_B = 8 + 5\ (6.0) + 5\ (1) = 43$$

$\therefore$ parcel B is in Category $C_4$

Stage III: Computation of State Probabilities

From an evaluation of empirical data, the elements in the category-impact state matrix below had been determined beforehand.

| | | IMPACT STATES | | | | |
|---|---|---|---|---|---|---|
| | | $R_1$ | $R_2$ | $R_3$ | $R_4$ | $R_5$ |
| CATEGORIES | $C_1$ | $n_{11}=2$ | $n_{12}=3$ | $n_{13}=4$ | $n_{14}=6$ | $n_{15}=10$ |
| | $C_2$ | $n_{21}=38$ | $n_{22}=5$ | $n_{23}=2$ | $n_{24}=3$ | $n_{25}=2$ |
| | $C_3$ | $n_{31}=5$ | $n_{32}=7$ | $n_{33}=8$ | $n_{34}=2$ | $n_{35}=3$ |
| | $C_4$ | $n_{41}=8$ | $n_{42}=12$ | $n_{43}=52$ | $n_{44}=3$ | $n_{45}=5$ |

Each element $n_{kj}$ gives the frequency with which we observed parcels in category $C_k$ being in state $R_j$.

$$\text{Then } P(R_j) = \left(\sum_{k=1}^{4} n_{kj}\right) \Big/ \left(\sum_{k=1}^{4}\sum_{j=1}^{5} n_{kj}\right)$$

$$= \left(\sum_{k=1}^{4} n_{jk}\right) \Big/ N$$

Where

$$N = \sum_{k=1}^{4}\sum_{j-1}^{5} n_{kj}$$

Therefore from Bayes' formula

$$P(R_j/C_k) = \frac{P(C_k/R_j)\ P(R_j)}{\sum_{j=1}^{5} P(C_k/R_j)\ P(R_j)}$$

$$= \frac{\left(n_{kj}/\sum_{k=1}^{4} n_{kj}\right)\ \sum_{k=1}^{4}\left(n_{kj}/N\right)}{\sum_{j=1}^{5}\left(n_{kj}/\sum_{k=1}^{4} n_{kj}\right)\left(\sum_{k=1}^{4} n_{kj}/N\right)}$$

$$= \frac{n_{kj}}{\sum_{j=1}^{5} n_{kj}}$$

Substituting the numerical values we get for parcel A.

$$P(R_1/C_2) = (38/50) = 76\%$$

$$P(R_2/C_2) = (5/50) = 10\%$$

$$P(R_3/C_2) = (2/50) = 4\%$$

$$P(R_4/C_2) = (3/50) = 6\%$$

$$P(R_5/C_2) = (2/50) = 4\%$$

Similarly for Parcel B

$$P(R_1/C_4) = (8/80) = 10\%$$

$$P(R_2/C_4) = (12/80) = 15\%$$

$$P(R_3/C_4) = (52/80) = 65\%$$

$$P(R_4/C_4) = (3/80) = 4\%$$

$$P(R_5/C_4) = (5/80) = 6\%$$

In addition to knowing the probabilities of parcels A and B being in the various impact states, one might wish to have a single estimate

of the impact state. Intuitively, we may choose the "most probable" state, the mode of the probability distribution. Another point estimator is the expected value. A more general choice of a point estimator is the Bayes estimation. This estimator $\hat{\theta}$ of the unknown impact state is defined with respect to a loss function $C(\theta, \hat{\theta})$ where $C(\theta, \hat{\theta})$ is the loss associated with predicting the impact state as $\hat{\theta}$ when it is in fact $\theta$. The Bayes estimator in this case is therefore that impact state which minimizes the expected loss for inaccurate predictions.

With various combinations of a land parcel descriptors (such as land use category, distance of parcel from improvement), one can generate a response surface which indicates the impact state that would occur under Bayes strategy. Such a response surface is shown on Figure 35.

It is evident from this simple example that the calculations leading to the various conditional probabilities would become very lengthy as the number of descriptors increases. The application of computers to aid in this approach is therefore much needed. Computers can be made to play many roles in the development of the concept just presented in this chapter. To list just a few, a computer can

(1) Produce a list of possible impact states for the land parcels in any given category.

(2) Compile statistics that relate a set of descriptors to impact states.

(3) Compute the probabilities of possible impact states.

(4) Resolve at high speeds the extensive information - storage and information - retrieval problems inherent in the maintenance of such large scale data collection as that called for in this chapter.

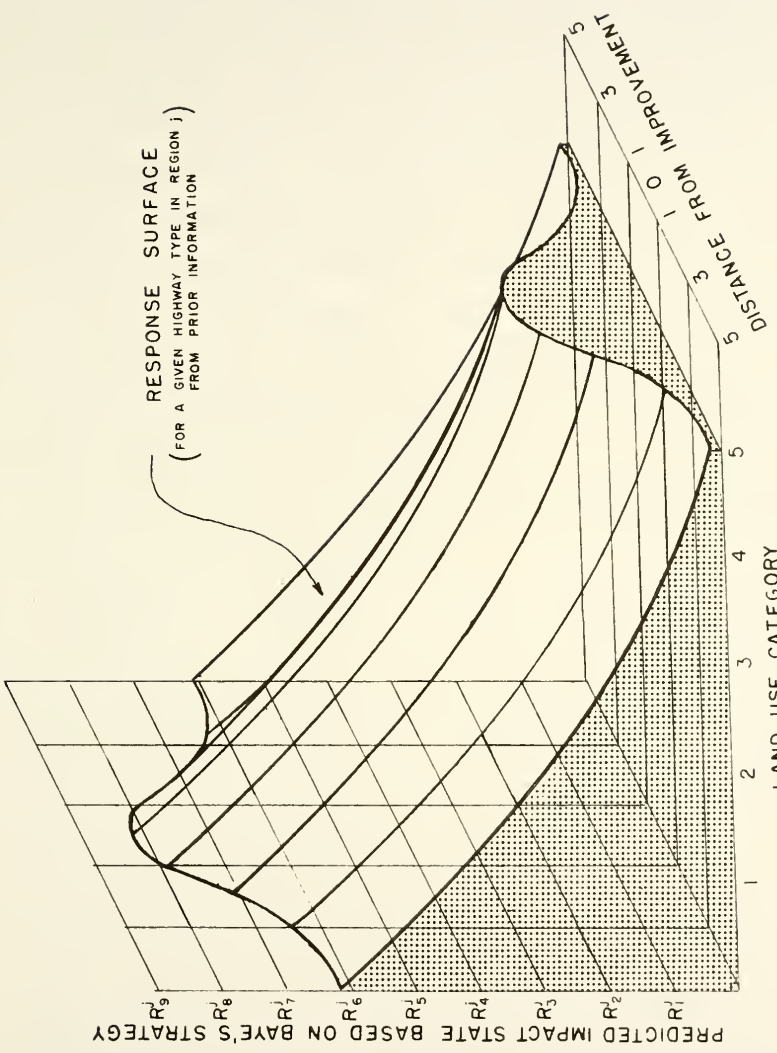



FIGURE35:RELATING DESCRIPTORS OF LAND PARCELS TO IMPACT STATES

(5) Permits easy updating of the posterior probabilities. In other words computers can continuously improve estimates of the conditional probabilities by making use of the most recent set of data. This prediction technique, therefore, becomes more accurate as more observations are fed into the data storage.

A computer however can do no more that it is programmed to do; it is an obedient servant. The successful application of computers is limited to the input data fed into it. Thus if a certain impact state-descriptor set (ISDS) relationship is not recorded in the computer memory no computations or statements can be made about it. Such empty cells in the ISDS matrix are analogous to cases in which a doctor observes some unfamiliar symptoms for which there is no previous knowledge in medical history. Thus the developed concept in this chapter can be applied for predictive purposes only within the range for which some historical information is available.

The cost of acquiring the needed data might be high if meaningful results using this approach are to accrue. An alternate technique to using past historical data might utilize the knowledge of a team of experts. Experience of these experts with land value impacts can form the basis for "degrees of belief" as to the probable states of impact that a given land parcel would encounter following a highway improvement. Each expert can provide a distribution function which conveys his measure of belief as to the relative likelihood of the various possible impact states. A problem then would be to reduce the various probability distributions obtained from the experts into a single distribution which represents consensus of the experts. This single resulting distribution would then

be used as an input to a formal Bayesian analysis.

The problem involved in developing such a single distribution from various probability distributions was discussed by Winkler ( 15 ) in his paper on "The Consensus of Subjective Probability Distributions".

# CHAPTER NINE

## CONCLUSIONS

Although the findings of this research are limited by the amount of data available, they provide a useful insight into changes in the value of land adjacent to and resulting from a highway improvement. Based on the preceding analyses, results and discussion, the following conclusions are summarized:

1. Variables defining the type of highway, the type of land use before the highway improvement, the type of access control between a land parcel and the highway improvement, the areal location of the parcel, the size of the parcel and the time elapsed after completion of the highway improvement explained to a substantial degree (slightly more than 87%) for Indiana parcels the variation in mean change of land value from before to after a highway improvement.
2. For the Indiana data, the variables denoting type of highway, areal location of the parcel and land use type were the most important in explaining the observed changes in land value following a highway improvement. Variables denoting type of access control and time elapsed after completion of the highway improvement were of lesser importance while size of parcel was found to be relatively unimportant.
3. An analysis of changes in mean effects of highway improvements for classes within each independent variable revealed the following trends:

A. Parcel along secondary and primary routes had greater increases in land value than parcels along Interstate routes (Only limited Indiana data available).

B. Both Indiana and Florida data revealed that parcels located in urban areas had a much greater increase in land value than parcels in rural areas.

C. Data from the two states revealed that residential parcels experienced greater increases in land value than parcels used for agricultural purposes. The Florida data also showed that parcels used for commercial purposes before the highway improvement suffered some loss in value. The Indiana data was too limited to be useful for any conclusion relative to commercial property.

D. Parcels with no or partial access control between them and the highway improvement had greater increases in land value than those with full control of access. (Again only Indiana data were available).

E. Data from both Indiana and Florida indicated that maximum increase in land value occurs during the eighteen to forty-one month period after the completion of the highway improvement.

F. Analyses of data from both States with respect to size of parcel revealed that small parcels (less than fifteen acres) had much greater increases in land values than larger parcels. Parcels larger than 50 acres in size suffered a decrease in mean land value.

4. Overwhelming evidence exists from the data, especially that from Indiana, for the presence of a functional relationship between the dependent and independent variables used in this research. The ratio of regression to residual mean squares was in all cases greater than the value that would be expected under the assumption that the independent variables have no systematic effects on the dependent variable.
5. The good statistical showing of models developed from the Indiana data suggests that with a stronger data base relatively simple models can be constructed to predict impact of highway improvements on land value.
6. There is abundant evidence in the data analyzed in this research that highway improvements enhance land values of parcels adjacent to them. Mean recovery rates of parcels in the states of Indiana and Florida were found to be about 640% and 400% respectively.
7. The dummy variable technique provided a useful tool for treating curvilinearity in the quantitative variables and for handling the discontinuous or categorical variables in the multiple regression analysis.
8. Of the three measures of land value impact used in this research, models using the change in land value expressed in constant dollars (DDIFF) as the dependent variable provided the highest coefficient of multiple determination. However one of the models using recovery rate (RVALUE) as the measure of impact produced the lowest coefficient of variation. In addition to

their high $R^2$ values, models using DDIFF involved less data preparation than those using RVALUE as the dependent variable.

9. The development of a conceptual framework for the non-deterministic approach outlined in Chapter Eight appears to provide an improved and realistic method for solving the complex and total highway impact problem.

# CHAPTER TEN

# RECOMMENDATIONS FOR EXTENSION AND FURTHER RESEARCH

The preceding analyses and results of this research suggested many areas in which a deeper insight into the problem of measuring highway impacts on land values is needed. Further research is therefore suggested in the following areas:

1. Sufficient data on the variables suggested by this research should be collected for the purpose of developing a reliable predictive model. The data for such research should also include measurements on other variables that have been found by other studies to be related to the land value impact problem.
2. The sensitivity of observed changes in land value to the density of highway improvements in a given area needs to be explored. It would appear that an inverse relationship exists between the magnitude of land value changes and the density of highway improvements in an area. This suggested hypothesis follows the same line of argument as Ricardo's Law of Diminishing Returns. This law asserts that as the units of a variable factor of production are applied to a fixed factor, additional output rises by lesser amounts. Bernoulli's utility concept provides extra evidence for this hypothesis. If this hypothesis is true for land value impacts it would provide some

light on why the same type of highway improvements in two urban areas might produce different degrees of impact.

3. A study should be made following the guidelines suggested in chapter eight to develop the proposed non-deterministic approach to the impact problem. In addition, the dynamic nature of highway impacts might be used in developing a stochastic approach to the impact problem. In such a development, it can be conceptualized that the occurrence of impact states qualifies as a Markovian process. As such one could study the sequence of impact states which a land parcel undergoes at different time periods following a highway improvement. Then it can be logically stated that given a category of land parcels at a particular state $S_i$ at some time t after the opening of the highway improvement, there is an associated probability ($P_{ij}$) that this category of parcels will next be in impact state $S_j$ at time (t + 1). Assuming that this transition is independent of the states of these parcels at some previous time periods, then the probability of this sequence of impact states is equal to the product of the probability of being in state $S_i$ at time t and the transition probability summed over all possible states of $S_i$. Further research is needed to investigate the usefulness of such an application of the Markovian process to the impact problem.

# LIST OF REFERENCES

# LIST OF REFERENCES

## CHAPTER ONE


1. United States Department of Commerce, Highways and Economic and Social Changes, Bureau of Public Roads, 1964, 1 pp.

2. Hay, William W., An Introduction to Transportation Engineering, John Wiley & Sons, Inc., New York, 1967.

3. de Neufville, Richard, "Toward a Comprehensive System Analysis for Transportation Planning: An Urban Example", Papers 1967: Eight Annual Meeting, Transportation Research Foundation, 111 pp.

4. United States Department of Commerce, Highways and Economic and Social Changes, Bureau of Public Roads, 1964, 5 pp.

5. Fisher, Joseph, "The Natural Environment", The Journal of the American Academy of Political and Social Science, Vol. 1, May 1967, pp. 127-140.

6. Stanford Research Institute, The Value of Time for Passenger Cars: Further Theory and Small Scale Behavioral Studies, Menlo Park, California, September, 1964, pp. 71-82.

7. Cherner, M., "Property Values as Affected by Highway Landscape Developments", Highway Research Record No. 53, Roadside Development, 1963 pp. 4-7.

8. Kemp, Barbara, H., "Social Impact of a Highway on an Urban Community", Highway Research Record No. 75, Highway Research Board, Washington, D. C.

9. Hennes, R. G., "Highways as an Instrument of Economic and Social Change", Highway Research Board Special Report 56, 135 pp.

10. Tybout, Richard A., "The Problem of Non-User Revenue", Highway Research Record No. 20.

11. Oregon State Highway Department, Economics of Highway Planning, Technical Bulletin No. 7, 1937.

LIST OF REFERENCES (continued)

12. AASHO, Road Users Benefit Analyses for Highway Improvements, Washington, D. C., 1960.

13. Haig, Robert M., "Major Economic Factors in Metropolitan Growth and Arrangement", Regional Survey of New York and its Environs Vol. 1, 1927, pp. 38-40.

14. House of Representatives Document No. 120 84th, Congress 1st Session 1955.

15. Zettel, Richard M., "Effect of Limited Access Highways on Property and Business Values", Paper presented at the 24th Annual Meeting of Institute of Traffic Engineers, Buffalo, New York, 1953.

16. Thiel, Floyd I., Yasnowsky, Jr., J., "Benefits of Highway Beautification", Public Roads, Vol. 35, No. 1, April 1968 3 pp.

17. Garrison, W. L., Berry, B. J. L., "A Source of Theory for Highway Impact Studies", Appendix E, Highway Research Board Special Report, 28, Washington, D. C., 1957.

18. Stover, V. G., "A Study of Remainder Parcels Resulting from the Acquisition of Highway Right-of-Way", Ph.D. dissertation, Purdue University, June 1963.

CHAPTER TWO

1. Goldstein, Sidney, "Economic Evidence in Right-of-Way Litigation", Public Roads, Vol. 32, June 1962.

2. Fulco, Roy J., "General Summary of Remainder Studies, Right-of-Way, Vol. 13, No. 6, December 1966.

3. Broderick, G. V., and Thiel F. I., "Highway Severance Damage Studies - Some General Findings", Highway Research Record No. 54, 1963.

4. Stover, V. G., "A Study of Remainder Parcels Resulting from the Acquisition of Highway Rights-of-Way", Ph.D. dissertation, Purdue University, June 1963.

5. Garrison, W. L., et. al., Studies of Highway Development and Geographical Change, University of Washington Press, 1959.

LIST OF REFERENCES (continued)

6. House Document No. 72, Final Report of the Highway Cost Allocation Study, (Part VI), 87th Congress 1st Session, (Washington: Government Printing Office, 1961).

7. Hensen, R. J., Mathias, J. S., Michael, H. L., "Impact of Lebanon Bypass 1950-1963, Purdue University, Indiana 1964.

8. Evans, E. G., "Impact of Kokomo Bypass 1950-64", Purdue University, Indiana 1966.

9. Michael, H. L., "An Economic Evaluation of the Kokomo Indiana By-pass", M. S. Thesis, Purdue University, August 1951, 3 pp.

10. Fletcher, J. A., Jr., "A Study of the Early Effects of a Rural Interstate Highway", M. S. Thesis, Purdue University, June 1961.

11. Lohr, A. F., "Early Impact of a Highway Improvement on an Urban Area", Joint Highway Research Project Report No. 14, Purdue University, May 1962.

12. Fleishman, Edward, R., "The Impact After Seven Years of a Highway Improvement in a Small City", M. S. Thesis, Purdue University, June 1968.

13. May, A. D., Jr., "An Evaluation of Certain Limited Access Highways", Ph.D. Thesis, Purdue University, January 1955.

14. Pinnell, Charles, "An Evaluation Study of Two Limited Access By-Passes in Indiana", M. S. Thesis, Purdue University, January 1958.

CHAPTER THREE

1. Adkins, William G., Effects of the Dallas Central Expressway on Land Value and Land Use, Bulletin 6, Texas Transportation Institute.

2. Kelly, John, F., Residences and Freeways, California Highways and Public Works, Vol. 36, No. 3-4, March-April 1957.

LIST OF REFERENCES (continued)

3. Netherton, Ross, D., Control of Highway Access, University of Wisconsin Press, Madison 1963 (351 p.).

4. "Washington Highway Economic Impact Studies", by R. G. Hennes, B. O. Wheeler and W. L. Garrison, HRB Bulletin 169, 1957 (pp. 60-74).

5. Garrison, W. L., et, al., Studies of Highway Development and Geographical Change, University of Washington Press, 1959.

6. Hensen, R. J., Mathias, J. S. and Michael, H. L., Impact of Lebanon Bypass, Purdue University, October 1964.

CHAPTER FOUR

1. Saaty, T. L., "Mathematical Models in Operations Research," McGraw Hill, 1959, Chapter 2.

2. Lowry, I. S., "A Short Course in Model Design," Journal of the ATP, XXXI, No. 2 (May 1965).

3. Duncan, Acheson J., Quality Control and Industrial Statistics, Homewood, Illinois, Richard D. Irwin Inc., 1959.

4. Harper, B. C. S. and Edwards, H. M., "A Study of the Generation of Person Trips by Areas in the Central Business District" Queens University, Kingston Ontario, Canada, 1960, 34 pp.

5. Tintner, Gerhard, Econometrics, New York, John Wiley & Sons, Inc., 1967, 85 pp.

6. Ezekiel, M. and Fox, K. A., Methods of Correlation and Regression Analysis, New York, Wiley and Sons, Inc., 1959.

7. Suits, D. B., "Use of Dummy Variables in Regression Equations", Journal of American Statistical Associations, Vol. 52, December, 1957, pp. 548-551.

8. Melichar, E., "Least Squares Analysis of Economic Survey Data", Proceedings of the Business and Economic Statistics Section, Washington, D. C., American Statical Association, 1965.

## CHAPTER FIVE


1. Biomedical Computer Programs, University of California, Los Angeles 1964.

2. Florida State Roads Department, Severance Study - Urban.

3. Florida State Road Department, Severance Study - Rural.

4. Stover, Vergil, G., "A Study of Remainder Parcels Resulting from the Acquisition of Highway Rights-of-Way", Ph.D. dissertation, Purdue University, June 1963.


## CHAPTER SEVEN


1. Draper, N. R., and Smith, H., Applied Regression Analysis, New York, John Wiley & Sons, Inc., 1966, 64 pp.

2. Wetz, J. M., "Criteria for Judging Adequacy of Estimation by an Approximating Response Function", A Ph.D. dissertation, University of Washington 1964.

3. Stover, V. G., "A Study of Remainder Parcels Resulting from the Acquisition of Highway Rights-of-Way", Ph.D. dissertation, Purdue University, June 1963.


## CHAPTER EIGHT


1. United States Department of Commerce, Highways and Economic and Changes, Bureau of Public Roads, 1964.

2. Hoffer, E., Computer Diagnosis: A Review and Discussion, Cambridge Scientific Center, May 1967.

3. Ledley, R. S. and Lusted, L. B., "The Use of Electronic Computers to Aid in Medical Diagnosis", Proceedings of the Institute of Radio Engineers, 47, (1959), 1970-1977.

4. Warner, H. R., "Medical Diagnosis Using a Digital Computer", Third IBM Medical Symposium, (1961)

5. Pavzen, Emmanuel, Modern Probability Theory and its Applications, John Wiley & Sons, Inc., 1960, New York 119 pp.

6. Biometrika, Vol. 46, 1958, (pp. 293-315).

LIST OF REFERENCES (continued)

7. Warner, Homer R., et. al., "A Mathematical Approach to Medical Diagnosis" JAMA 177, 177-183, July 22, 1961.

8. Overall, J. E. and Williams, C. M., "Conditional Probability Program for Diagnosis of Thyroid Function", JAMA 183: 307-313, February 2, 1963.

9. Fitzgerald, L. T., et al., "Modified Program for Computer Diagnosis of Thyroid Disease", Radiology 82 : 334, February, 1964.

10. Dubois, Phillip H., Multivariate Correlation Analysis, New York, Harper and Brothers, 1957.

11. Wortman, Robert H., "A Multivariate Analysis of Vehicular Speeds on Four-Lane Highways", University of Illinois, June 1963.

12. Harman, Harry, H., Modern Factor Analysis, University of Chicago Press, Second Edition, 1967.

13. Fisher, R. A., "The Use of Multiple Measurements in Laxonomic Problems," Annals of Engenics, Vol. 7 (1936), 179 pp.

14. Casetti, Emilio, Multiple Discriminant Functions, Technical Report No. 11 of ONR Task No. 389-134, Office of Naval Research Geography Branch, Northwestern University, Evanston, Illinois, May, 1964.

15. Winkler, Robert, L., "The Consensus of Subjective Probability Distributions", Management Science Application (U. S.) 15 (1968) 2 (Oct.) pp. B-61-B-75.

General References


1. Li, Jerome C. R., Statistical Inference II, (Chapter 33), Edward Brother, Inc., Ann Arbor, Michigan 1964.

2. Tomek, William G., "Using Zero-One Variables with Time Series Data in Regression Equations," Journal of Farm Economics, 45 1963 1963 pp. 814-822.

3. Davies, M., "Multiple Linear Regression Analysis with Adjustment for Class Differences", Journal of the American Statistical Association, 56, 1961 pp. 729-735.

4. Draper, N. R., Smith, H., Applied Regression Analysis, John Wiley & Sons, Inc., 1966, pp. 134-141.

5. Ostle, B., Statistics in Research, Second Edition, Iowa State University Press, Ames, Iowa, U.S.A., 1963.

6. Thrall, R. M., Foundations and Tools of Operations Research and the Management Sciences, University of Michigan, Summer Conference Notes 1962.

7. Saaty, T. L., "Mathematical Models in Operations Research", McGraw Hill, 1959, Chapter 2.

8. Lowry, I. S., "A Short Course in Model Design", Journal of the AIP XXXI, No. 2, May 1965.

9. Warner, Homer, R., et. al., "A Mathematical Approach to Medical Diagnosis", JAMA 177: 177-183, July 22, 1961.

10. Overall, J. E., and Williams, C. M., "Conditional Probability Program for Diagnosis of Thyroid Function", JAMA 183, 307-313, February 2, 1963.

11. Fitzgerald, L. T., et. al., "Modified Program for Computer Diagnosis of Thyroid Disease", Radiology 82: 334, February 1964.

12. Tsao, R. F., A Second Order Exponential Model for Multidimensional Dichotomous Contingency Tables with Applications in Medical Diagnosis, August, 1967, IBM, Data Processing Division, Cambridge Scientific Center.

13. Burnbaum, A. and Maxwell, A. E., "Classification Procedures Based on Baye's Formula", Applied Statistics, (London), 1960, Vol. 9, No. 3, 152 pp.

General References (continued)


14. Ledley, R. S. and Lusted, L. B., "The Use of Electronic Computers to Aid in Medical Diagnosis" Proceedings of the Institute of Radio Engineers, (New York), 47 1959, 1970-1977.

15. Suits, D. B., "Use of Dummy Variables in Regression Equations" Journal of American Statistical Association, Vol. 52, December 1957, pp. 548-51.

16. Morgan, J. N., "A Note on the Interpretation of Multiple Regression, Using Dummy Variables", Unpublished Notes Survey Research Center, Institute for Social Research, University of Michigan, April 28, 1964.

17. Guidelines for Trip Generation Analysis, Appendix D, "Multiple Regression Using Dummy Variables", pp. 99-108, by BPR, June 1967.

18. Melichar, E., 'Least Squares Analysis of Economic Survey Data", Proceedings of the Business and Economic Statistics Section, Washington, D. C., American Statistical Association 1965.

19. "Evaluation of Economic Impact", by A. S. Lang and Martin Wohl, HRB Bull. 268, 1960, pp. 105-119.

20. "Reflections on Concepts for Impact Research", by R. H. Stroup and L. A. Vaugha, HRB Bull., 311, 1962, pp. 1-12.

21. "Guide for Highway Impact Studies", U. S. Bureau of Public Roads, Division of Highway and Land Administration.

22. "Economics of Highway Location", A Critique of Collateral Effect Analysis", by Professor W. A. Pillsbury, HRB Bull, 75, 1965.

23. "Land Economic Studies in Indiana", by V. G. Stover and H. L. Michael, Proceedings of the 49th Annual Road School, Purdue University, September 1963.

24. "AASHO Road User Benefit Analyses for Highway Improvements," AASHO Committee on Planning and Design Policies, Washington, D. C., 1952.

25. U. S. Congress, 87th, 1st Session, Final Report on the Highway Cost Allocation Study, House Documents 54 and 72, 1961, Washington, D. C. U. S. Governtment Printing Office.

26. "Land Values and the Measurement of Highway Benefits", by H. Mohring, Journal Political Econ., June, 1961.

27. "Studies of Highway Development and Geographic Change," by W. L. Garrison, University of Washington Press, Seattle.

General References (continued)


28. Pendleton, W. C., "Relation of Highway Accessibility to Urban Real Estate Values", HRR No. 16, Highway Research Board, Washington, D. C.

29. Cribinns, P. D., Hill, W. T. and Seagraves, H. O., "Economic Impact of Selected Sections of Interstate Routes on Land Values and Land Use", HRR No. 75.

30. Wendt, P. F., "Influence of Transportation Changes on Urban Land Uses and Values", HRB. Bull. 268.

31. Garrison, W. L., "Statistical Analysis of a Road Improvement Impact Problem", HRB Special Report 28, 1957.

32. McRain, W. C., "Economic and Social Impact of the Connecticut Turnpike", HRB Bull. 189.

33. Garrison, W. L., "Approaches to Three Highway Impact Problems", HRB Bull. 227.

34. Broderick, George V., and Thiel, Floyd I., "Highway Severance Damage Studies - Some General Findings, HRR No. 54, Land Acquisition, 1963.

35. Adkins, W. G., "Land Values Impacts of Expressways in Dallas, Houston and San Antonio, Texas", HRB Bull. 227.

36. Bone, A. J., and Wohl, Martin, "Massachusetts Route 128 Impact Study", HRB Bull. 227, Highway Research Board, Washington, D. C.

37. Fletcher, J. A., Jr., "Studies of Highway Impact in Indiana - Early Effects of a Portion of Interstate 65", (Unpublished report) Joint Highway Research Project, Purdue University, June 1961.

38. McGilly, F. J., "Determining the Impact of Expressways, Better Roads, Vol. 30, No. 12, December 1960, pp. 31-33.

39. McCoy, George T., "Techniques of Making Land Economic Studies", American Highways, Vol. 35. No. 3, July 1956.

40. Mohring, H. D and Harwitz, M, "Highway Benefits", Northwestern University Press, 1962.

41. "Indiana's Interstate Highway System", A Highway Transportation Research Report, Transportation Department, Indiana State Chamber of Commerce, Indianapolis, Indiana, January 1960.

42. "Land Values Before and After Freeway Development by Rudolph", Hess American City, Vol. 57, No. 10 October 1952.

General References (continued)


43. "A 15 Year Study of Land Values and Land Use Along the Gulf Freeway Houston, Texas", Norris and Elder Consulting Engineers, Houston, Texas, 1956, pp. 146-149.

44. Adkins, William G., "Effects of the Dallas Central Expressway on Land Values and Land Use", Texas Transportation Institute Bulletin 6., September, 1957.

45. Longley, J. W. and Goley, B. T., "A Statistical Evaluation of the Influence of Highways on Rural Land Values in the United States , HRB 327, pp. 21-55.

46. 'Recreational Impact of Multi-Purpose Reservoirs, Judson S. Matthias Ph.D. Thesis, 1967.

47. "IBM Coding Procedure for Highway Impact Studies", Economic Studies Section, Planning Division, Oklahoma State Highway Department, Oklahoma City, Oklahoma, June 1960.

48. Cornell, Robert C., "Economic Impact Studies and Their Applications", Planning Conference Procedure, Western Association of State Highway Official, Phoenix, Arizona, April, 1958, pp. 45-48.

49. Goldstein, Sidney, "Non-User Benefits from Highway", HRB, HRR No. 20, 1963.

50. Trier, Robert, J., and Kubitz, V. S., "The Impact of Modern Highways on American Indian Country", HRR No. 149, p. 56.

51. Clair, G. P. and Others, "The Measurement of Vehicular Benefits", HRR No. 138, 1966.

52. Gamble, H. B., Raphael, D. L., and Sawerlander, O. H., "Direct and Indirect Economic Impacts of Highway Interchange Development", HRR No. 149, 1966, 42 pp.

53. Lang, A. S., and Roberts, P. O., "An evaluation of Techniques for Highway User Cost Compulation", HRB No. 320.

54. Thiel, Floyd I., "Benefits of Highway Beautification", Public Roads, April 1968, Vol. 35, No. 1 pp. 1-8.

55. Goldstein, Sidney, "Economic Evidence of Right-of-Way Litigation", Public Roads, June 1962, Vol. 32, No. 2.

56. Thiel, Floyd I., "Severance Case Studies-Bridging the Gap Between Findings and Their Application", Public Roads, April 1967, Vol. 34, No. 7.

General References (continued)

57. Lum, Y. T., "More Explanation of the Assumptions, Doctrines and Implication of Right-of-Way Valuations - The Meaning, Comparison and Application of Market Data", Right-of-Way, August 1964, Vol. 11, No. 4.

58. Grace, Arthur B., Jr., "Severance Damage Studies", Right-of-Way, February 1966, Vol. 13, No. 1.

59. McGough, Bobby C., "A Treatise on Highway Impact Studies", Right-of-Way", June 1968, Vol. 15, No. 3.

60. Dilmore, Gene, "Land Value of Improved Properties", Right-of-Way, April 1967, Vol. 14, No. 2.

61. Fulco, Roy J., "General Summary of Remainder Studies, Right-of-Way, SRA, December 1966, Vol. 13, No. 6.

62. McGough, B. C., "Impact of Highway Improvements on Land Values as Observed in Two Florida Studies", Right-of-Way, December, 1965, Vol. 12, No. 6.

63. Ryan, F. E., "A Method of Measuring Changes in the Value of Residential Properties", Land Acquisition 1959, HRB Bulletin 232.

64. "Changes in Westchester and How People Feel About Them", Westchester County Department of Planning, White Plains, New York, 1955, pp. 121-

65. Kemp, Barbara, H., "Social Impact of a Highway on an Urban Community", Highway Research Record No. 75, HRB, Washington, D. C.

66. "Classification Procedures Based on Baye's Formula", Applied Statistics, London 1960, Vol. 9, No. 3, 152 pp.

67. Dietz, Stephen K., "Significance Test for Accident Reductions Based on Classical Statistics and Economic Consequences", Transportation Science, Vol. 1, No. 3, August 1967, pp. 200-217.

69. Warner, H. R., Toronto, A. F., Veasy, L. G. and Stephenson, R. A., "Mathematical Approach to Medical Diagnosis", Journal of the American Medical Association, (Chicago) 177, 1961.

# VITA

# VITA

Edward Irogueni Isibor was born June 9, 1940 in Benin City, Nigeria to Mr. Joseph Ariebuwa Isibor and Mrs. Agnes Aghaku Isibor. He attended Lagos Government School and St. Peter's School, Benin City for his elementary education. For his high school education, he attended Edo College, Benin City from 1954-1959 and Government College, Ibadan from 1960-1961. He taught at the Immaculate Conception College for most of 1962.

He was awarded a scholarship for his undergraduate work at Howard University, Washington, D. C. by the African Scholarship Program for American Universities. Among his many academic honors at Howard University, was the faculty award for "the most outstanding student" in civil engineering. He was also cited for distinction in "Who's Who Among Students in American Universities and Colleges" Publication 1965, and in "Africa in the United States" edited by Professor Vernon McKay of the John Hopkins University.

He was appointed a Graduate Research Assistant at the Massachusetts Institute of Technology while he was working for his Master of Science Degree in Civil Engineering.

After graduating from M.I.T. in January 1967, he was employed by Edward and Kelcey Consulting Firm at their Boston office.

He has been employed as a Research Assistant by Purdue University since September 1967 and has been working on the Ph.D. Degree since that time.

VITA (continued)

He is an associate member of the American Society of Civil Engineers, a student member of the American Statistical Association, Institute of Traffic Engineers, a Supporting Member of the Highway Research Board and a Student Associate Member of the Operation Research Society of America. He is a member of Tau Beta Pi.

He is married and has a son.